

---

A Guide to Using

# SEER\*Stat

Version 3.0

National Cancer Institute  
Cancer Statistics Branch, DCCPS  
EPN 343J; MSC 7352  
6130 Executive Blvd.  
Bethesda, Maryland 20892-7352



# Table of Contents

Table of Contents .....	3
Introduction to SEER*Stat .....	1
<b>About SEER*Stat</b> .....	1
<b>Statistics - Frequency/Rate Sessions</b> .....	1
<b>Statistics - Survival Sessions</b> .....	1
<b>National Cancer Institute</b> .....	2
<b>SEER Cancer Data</b> .....	2
Getting Started .....	3
<b>What You Will Need</b> .....	3
<b>Overview</b> .....	3
<b>Preferences</b> .....	4
<b>Remote Menu</b> .....	5
<b>Using Your Own Data</b> .....	5
Getting Help .....	7
<b>Help Contents</b> .....	7
<b>Context-Sensitive Help</b> .....	7
<b>Technical Support</b> .....	8
Frequency/Rate Sessions .....	9
<b>Open a Frequency/Rate Session</b> .....	9
<b>Frequency/Rate Session Overview</b> .....	9
<b>Save a Frequency/Rate Session</b> .....	10
<b>Frequency/Rate Data Tab</b> .....	10
Database for a Frequency/Rate Session .....	11
Statistic for a Frequency/Rate Session .....	12
Statistics Requiring a Population File .....	12
Parameters for a Frequency/Rate Session .....	13
<b>P-Values for a Frequency/Rate Session</b> .....	13
Edit P-Values Used in a Frequency/Rate Session .....	13
Standard Million for Frequency/Rate Session .....	14
Age Variable for a Frequency/Rate Session .....	14
Trend Variable for a Frequency/Rate Session .....	14
<b>Frequency/Rate Selection Tab</b> .....	15
Age at Diagnosis Selection for a Frequency/Rate Session .....	15
Race, Sex, Year Dx, Registry, County Selection for a Frequency/Rate Session .....	16
Other Case Variables Selection for a Frequency/Rate Session .....	16
<b>Frequency/Rate Table Tab</b> .....	16
Add a Table Variable to a Frequency/Rate Session .....	17
Order the Table Variables in a Frequency/Rate Session .....	18
Remove a Table Variable from a Frequency/Rate Session .....	18
<b>Frequency/Rate Output Tab</b> .....	18

Title for a Frequency/Rate Matrix .....	19
Percentages for Frequency/Rate Sessions .....	19
Hide Statistics When Less Than ____ Cases .....	20
Display Standard Errors and Confidence Intervals.....	20
Display p-Values for EAPC Significance Testing .....	21
Display Rates as Cases Per ____.....	21
<b>Survival Sessions.....</b>	<b>23</b>
<b>Open a Survival Session .....</b>	<b>23</b>
<b>Survival Session Overview.....</b>	<b>23</b>
<b>Save a Survival Session .....</b>	<b>24</b>
<b>Survival Data Tab .....</b>	<b>24</b>
Database for a Survival Session .....	25
Survival Rates Definitions .....	25
Observed and Relative Survival .....	26
Observed Survival Only.....	26
Expected Rate Table.....	27
<b>Survival Selection Tab.....</b>	<b>28</b>
Standard Survival Selections .....	28
<i>Select Only Microscopically Confirmed .....</i>	<i>28</i>
<i>Select Only Actively Followed .....</i>	<i>29</i>
<i>Select Only Malignant Behavior .....</i>	<i>29</i>
<i>Exclude All Second and Later Primaries .....</i>	<i>29</i>
<i>Exclude All Death Certificate and Autopsy.....</i>	<i>29</i>
Expected Rate Table Standard Survival Selections .....	29
<i>Exclude All Unknown Races .....</i>	<i>29</i>
<i>Exclude All Unknown Ages .....</i>	<i>30</i>
<i>Exclude All Sex not Male or Female .....</i>	<i>30</i>
<i>Exclude All Invalid Age Dates .....</i>	<i>30</i>
Case Selection for a Survival Session .....	30
<b>Survival Table Tab .....</b>	<b>31</b>
Add a Page Variable to a Survival Session.....	31
Order the Page Variables in a Survival Session.....	32
Remove a Page Variable from a Survival Session.....	32
Avoid Inadvertent Exclusions .....	32
<b>Survival Parameters Tab .....</b>	<b>32</b>
Dates for a Survival Session .....	33
<i>Begin Date .....</i>	<i>33</i>
<i>End Date .....</i>	<i>34</i>
<i>Study Cutoff Date.....</i>	<i>34</i>
<i>Lost to Follow Up Date.....</i>	<i>34</i>
Intervals for a Survival Session .....	35
<i>Number of Intervals.....</i>	<i>35</i>
<i>Months Per Interval.....</i>	<i>35</i>
<i>Special Intervals.....</i>	<i>35</i>
<i>Z-Test Interval.....</i>	<i>36</i>
Vital Status for a Survival Session .....	36

<b>Survival Output Tab</b>	<b>36</b>
Title for a Survival Matrix	37
Display Tables	37
Standard Life Table	38
Cumulative Summary	38
Flag Cum Relative SE's Greater Than ____ %	38
Suppress Pages With Less Than ____ Cases	39
Adjust Relative Rates over 1.0	39
Adjust Increasing Relative Rates	39
Adjust for Heterogeneity in Withdrawal	39
<b>Case Listing Sessions</b>	<b>41</b>
<b>Open a Case Listing Session</b>	<b>41</b>
<b>Case Listing Session Overview</b>	<b>41</b>
<b>Save a Case Listing Session</b>	<b>42</b>
<b>Case Listing Data Tab</b>	<b>42</b>
Database for a Case Listing Session	42
<b>Case Listing Selection Tab</b>	<b>42</b>
<b>Case Listing Table Tab</b>	<b>43</b>
Add a Table Variable to a Case Listing Session	43
Order the Table Variables in a Case Listing Session	44
Remove a Table Variable from a Case Listing Session	44
<b>Case Listing Output Tab</b>	<b>44</b>
Title for a Case Listing Matrix	44
<b>Dictionary</b>	<b>45</b>
Open the Dictionary	45
Dictionary Window	45
Import User Variables	47
Import User Variable Definitions Window	47
Export User Variables	48
Export User Variable Definitions Window	48
<b>Database Variables</b>	<b>49</b>
Standard Variables	50
User-Defined Variables	50
Merged Variables	51
<b>Working With Variables</b>	<b>51</b>
Create a User-Defined Variable	51
Create a Merged Variable	52
Modify a User-Defined Variable	52
Modify a Merged Variable	53
Delete User-Defined or Merged Variables	53
Save a User-Defined Variable to the Database Dictionary	53
Save a Merged Variable to the Database Dictionary	54
Edit Variable Window	54
Open the Edit Variable Window	54
Edit Variable Window	54

Add a New Grouping to a User-Defined Variable .....	56
Add All Variable Values to a User-Defined Variable.....	56
Add Rest of the Variable Values to a User-Defined Variable .....	57
Update the Definition of an Existing Grouping in a User-Defined Variable....	57
Rename an Existing Grouping in a User-Defined Variable.....	58
Order the Groupings of a User-Defined Variable .....	58
Delete an Existing Grouping from a User-Defined Variable .....	58
Merged Variable Window .....	59
Open the Merged Variable Window.....	59
Merged Variable Window .....	59
Add a New Grouping to a Merged Variable.....	61
Add All Groupings to a Merged Variable .....	62
Add All Merged Groupings Window.....	63
Edit a Grouping Definition in a Merged Variable .....	63
Delete a Grouping from a Merged Variable.....	64
Rename a Grouping in a Merged Variable.....	64
Order the Groupings of a Merged Variable .....	64
Include “Other” Grouping in a Merged Variable.....	65
<b>Working with Search Statements.....</b>	<b>67</b>
<b>Search Statements.....</b>	<b>67</b>
Selection Window.....	67
Accessing the Selection Window.....	69
Building Search Statements with Search Lines .....	69
Create a Search Line .....	69
Edit a Search Line.....	70
Find Window .....	70
Selection(s) Box on the Selection Window .....	70
Select Groupings for a Search Line.....	70
Select Continuous Values for a Search Line .....	71
Select Both Continuous Values and Groupings for a Search Line .....	71
<b>Linking Search Lines to Make a Search Statement.....</b>	<b>72</b>
Conjunctions for Search Statements .....	72
Precedence in a Search Statement.....	72
Link Search Lines With a Conjunction (And/ Or) .....	73
Add Parentheses to a Search Statement .....	74
Add a “Not” to a Search Statement .....	74
<b>Edit the Search Statement.....</b>	<b>75</b>
Edit a Search Line or Change the Conjunction .....	75
<b>Add a Search Line.....</b>	<b>75</b>
Edit Parentheses in a Search Statement.....	76
Edit “Not” in the Search Statement.....	76
Delete a Search Line from a Search Statement .....	77
Order Search Lines in a Search Statement.....	77
<b>Tips For Accurate Search Statements.....</b>	<b>78</b>
Select Variables Systematically .....	78

Create an Exact Search Statement .....	79
Avoid Variable Groupings that Overlap .....	79
<b>Execute a Session.....</b>	<b>80</b>
<b><i>Job Status Window .....</i></b>	<b><i>80</i></b>
<b><i>Finish Remote .....</i></b>	<b><i>80</i></b>
<b><i>Cancel a Job.....</i></b>	<b><i>81</i></b>
<b><i>Set the Job Priority .....</i></b>	<b><i>81</i></b>
<b>Output Matrix.....</b>	<b>82</b>
<b><i>Open a Matrix .....</i></b>	<b><i>82</i></b>
<b><i>Save a Matrix .....</i></b>	<b><i>82</i></b>
<b><i>View the Session Which Produced the Matrix.....</i></b>	<b><i>82</i></b>
<b><i>Matrix Properties Window .....</i></b>	<b><i>83</i></b>
<b><i>Frequency/Rate Matrix.....</i></b>	<b><i>84</i></b>
Frequency/Rate Matrix Window .....	84
Frequency/Rate Matrix Options.....	85
Matrix Footnotes .....	86
<b><i>Frequency/Rate Matrix Navigation .....</i></b>	<b><i>86</i></b>
Page Through a Matrix.....	86
Select Columns or Rows in a Matrix.....	87
Lock Feature for a Matrix .....	87
Resize Columns or Rows in a Matrix.....	87
Reorder Column or Row Variables in a Frequency/Rate Matrix .....	88
Order Labels Window .....	88
<b><i>Survival Matrix.....</i></b>	<b><i>88</i></b>
Survival Matrix Window .....	88
Survival Matrix Options .....	89
Matrix Footnotes .....	90
<b><i>Survival Matrix Navigation .....</i></b>	<b><i>90</i></b>
Page Through a Matrix.....	90
Select Columns or Rows in a Matrix.....	91
Lock Feature for a Matrix .....	91
Resize Columns or Rows in a Matrix.....	91
<b><i>Case Listing Matrix .....</i></b>	<b><i>92</i></b>
Case Listing Matrix Window .....	92
Case Listing Matrix Options .....	93
<b><i>Case Listing Matrix Navigation .....</i></b>	<b><i>94</i></b>
Select Columns or Rows in a Matrix.....	94
Change Alignment of Values in a Matrix .....	94
Lock Feature for a Case Listing Matrix.....	95
Hide Columns in a Case Listing Matrix.....	95
Change Format of Variable Values in a Matrix.....	95
Resize Columns or Rows in a Case Listing Matrix.....	96
Reorder Column Variables in a Case Listing Matrix .....	96
<b><i>Change Matrix Font.....</i></b>	<b><i>96</i></b>
Font Window .....	97

Font Type .....	97
Font Style .....	97
Font Size .....	97
Font Sample .....	97
<b>Number Filter .....</b>	<b>98</b>
Accessing the Number Filter .....	98
Filter Window .....	98
Suppress if Zero Filter .....	99
Thousands Separator Filter .....	99
Leading Zero Filter .....	100
Fixed Width Filter .....	100
Rounding Filter .....	101
Negatives Filter .....	101
Prefix Characters Filter .....	101
Suffix Characters Filter .....	101
<b>Sharing Your Results with Other Users and Applications .....</b>	<b>102</b>
<b><i>Copy a Matrix to the Windows Clipboard .....</i></b>	<b><i>102</i></b>
<b><i>Print a Session or Matrix .....</i></b>	<b><i>102</i></b>
<b><i>Use Shading When Printing .....</i></b>	<b><i>103</i></b>
<b><i>Print User-Defined Variable Definitions .....</i></b>	<b><i>103</i></b>
<b><i>Exporting Results to Other Applications .....</i></b>	<b><i>103</i></b>
Export Data Files .....	104
Export Data Dictionary .....	104
Export Window .....	105
Output Variables as Numeric or Text in Export File .....	109
Line Delimiter in Export File .....	110
Field Delimiter in Export File .....	110
Missing Character in Export File .....	111
Enclose Fields Containing Delimiter in Quotes in Export File .....	111
Remove all Thousands Separators (Commas) from Export File .....	112
Remove Flags (Footnote Characters) from Export File .....	112
Output Variable Names Before Data .....	112
Edit Export Options .....	112
<b>Algorithms .....</b>	<b>114</b>
<b><i>Frequency/Rate Algorithms .....</i></b>	<b><i>114</i></b>
Algorithms for Rates .....	114
Algorithms for Trends .....	116
<b><i>Survival Algorithms .....</i></b>	<b><i>117</i></b>
Symbol Definitions for Survival Algorithms .....	118
Algorithms for the Survival Output Rate Table .....	118
Algorithms for the Special Survival Follow up Interval .....	120
Expected Survival Rates Table .....	120
Indexing the Expected Probabilities for Each Individual Patient .....	121
<b>Bibliography .....</b>	<b>122</b>



<b><i>Frequency/Rate Bibliography .....</i></b>	<b><i>122</i></b>
<b><i>Survival Bibliography .....</i></b>	<b><i>122</i></b>
<b><i>Index.....</i></b>	<b><i>124</i></b>



# Introduction to SEER\*Stat

## About SEER\*Stat

SEER\*Stat is a statistical package created for the analysis of SEER and other cancer databases. It was developed by Information Management Services, Inc. in consultation with the SEER Program of the National Cancer Institute. SEER\*Stat is designed to be an efficient and flexible tool with which to produce cancer frequencies, rates, survival statistics, and case listings. This system is a 32-bit Microsoft Windows application which provides the functionality and intuitive interface commonly found in the Windows environment.

## Statistics - Frequency/Rate Sessions

From a SEER\*Stat Frequency/Rate Session, you can calculate the following statistics for a population based database. If no population file is associated with the database, only frequencies and row/column percentages may be calculated. To calculate trends, a year variable (year of diagnosis or year of death) is also required.

- Frequencies with or without row or column percentages
- Crude rates (non-adjusted) with standard errors and confidence intervals
- Age-adjusted rates with standard errors and confidence intervals
- Trends over time as percent change (PC), from crude or age-adjusted rates
- Trends over time as Estimated Annual Percent Change (EAPC), from crude or age-adjusted rates, with confidence intervals
- Comparison of EAPC with zero or a comparison of two EAPCs with one another

## Statistics - Survival Sessions

From a SEER\*Stat Survival Session, you can calculate the following statistics for an incidence database with follow up information.

- Observed survival statistics
- Expected survival statistics and relative survival statistics
- Comparison of two survival curves (Z-statistic)
- Conditional survival (special intervals)

# **National Cancer Institute**

One of the major roles of the National Cancer Institute, which commissioned SEER\*Stat, is to collect, analyze and disseminate data useful in the prevention, diagnosis, and treatment of cancer. In support of this mission, the National Cancer Institute collects information on cancer incidence and survival through the Surveillance, Epidemiology and End Results (SEER) program.

The SEER program is composed of a set of population-based cancer registries across the United States which collect cancer data on a routine basis. The cancer data collected is submitted to the National Cancer Institute semi-annually for analysis.

For more information about the SEER program, please refer to the SEER Web page:

[www.seer.cancer.gov](http://www.seer.cancer.gov)

## **SEER Cancer Data**

The data delivered with SEER\*Stat originates from the National Cancer Institute's Surveillance, Epidemiology and End Results (SEER) program. The SEER program is a collection of population-based cancer registries in the United States which collect and submit cancer incidence and follow up data to the National Cancer Institute.

The databases supplied with SEER\*Stat are the SEER Public-Use database which are periodically distributed by the National Cancer Institute. They contain a select number of variables collected by the SEER program or generated by the National Cancer Institute. Two forms of the Public-Use data are delivered with SEER\*Stat. One is in a binary format used by SEER\*Stat and should not be accessed or viewed in a text editor. The other is in an ASCII format. The layout and data dictionary of the ASCII files are available on the CD.

# Getting Started

## What You Will Need

To use SEER\*Stat, you will need:

- 486- or Pentium-based personal computer (Pentium recommended)
- Microsoft Windows 95/NT
- 16 MB application RAM
- 2 MB hard disk space

Additional information on some SEER variables are available on the Internet. To link to this additional documentation, you must have an Internet browser installed and be connected to the Internet.

## Overview

SEER\*Stat allows you a great deal of freedom to request exactly the cancer statistics/values you want for your analysis.

### **Part 1: Session**

The analysis is set up in the session window. Each session consists of tabs on which you select the database subset, statistics, and appearance of your output matrix. You should work through each tab in order from left to right and from top to bottom to ensure that all options have been considered. However, changes can be made in any order. It is possible to work on multiple sessions simultaneously.

### **Part 2: Job**

Once the session is set up, you are ready to execute it as a job. While the job is executing, you can change the session or begin a new one without affecting the original job. It is possible to execute more than one job at a time.

### **Part 3: Matrix**

When the job has finished executing, the output matrix you requested is displayed. You can change the appearance of the output matrix, print it, copy it to the Windows clipboard, and/or export the statistics/values so they may be used in another application.

# Preferences

Access the Preferences Window by selecting Preferences from the File menu. Make changes to the fields as necessary. When you are finished, click OK.

## **Primary Data Location**

You may enter the directory path for the SEER\*Stat databases. If you do not know the entire path name, click the Browse button to select the path. This field is disabled if any session windows are open.

## **Secondary Data Location**

You may create your own SEER\*Stat database using SEER\*Prep. When a database is created by SEER\*Prep, this field is updated to the directory path of the new database. The directory path can also be entered manually. If you do not know the entire path name, click the Browse button to select the path. This field is disabled if any session windows are open.

## **User-Defined Variables Location**

You may enter the directory path where you will save user-defined variables. This will also be the directory from which the Save and Open functions will begin by default. If you do not know the entire path name, click the Browse button to select the path.

## **Default Priority**

Select high, normal, or low for the beginning priority setting for every new job. To read more about setting the priority for individual jobs refer to Set the Job Priority.

## **Use Shading When Printing**

The shading when printing option adds shading to the column and row titles of the printed output matrix. If you plan to make copies or FAX the matrix, it may be wise to turn off this option.

## **Save to Dictionary as Default**

New user-defined and merged variables may be created for use within a single session, or may be saved to the database dictionary for use in any session. When selected, the Save to Dictionary option at the bottom of the Edit Variable Window and Merged Variable Window will be checked by default for all new user-defined and merged variables.

## **Print User-defined Variable Definitions**

The Print User-defined Variable Definitions option allows you to choose whether or not to print the definitions of all user-defined variables used in a session. If selected, the title of the user-defined variable, the standard

variable on which it is based, the groupings, and all values associated with each grouping are listed on the session print out.

### **Default for Displaying Non-Survival Rates**

You may change the default Cases Per value shown on the Frequency/Rate Output Tab.

### **Default Font**

To change the default font of the output matrices, click the Change button. The Font Window will appear. When you have finished making your choices, click OK to return to the Preferences Window.

## **Remote Menu**

The Remote menu can be accessed from the File menu when you are utilizing SEER\*Stat within the client-server configuration. Once you have logged in, use the Remote menu to change your password or edit your user information.

### **Change Password**

To change your password:

- 1 Log into SEER\*Stat using your previous or assigned username and password.
- 2 Select Change Password from the Remote option on the File menu.
- 3 Enter your old password, then your new password twice in the designated text boxes and click OK.

### **Edit User Information**

This is where you can change your e-mail address. When you choose the Finish Remote option from the Job Status Window, notification of the completion of a SEER\*Stat execution is sent to this address.

## **Using Your Own Data**

You can use your own incidence and mortality data with populations and expected rates in SEER\*Stat. To do this you must first prepare your data for SEER\*Stat with a software package called SEER\*Prep. It is also possible to prepare your own expected rate tables using SEER\*Prep. SEER\*Prep is available on the SEER Program's web site under the topic of Scientific Systems:

[www.seer.cancer.gov](http://www.seer.cancer.gov)

When preparing your data using SEER\*Prep, keep in mind how you plan to utilize it within SEER\*Stat.

**Frequencies** can be calculated on any data prepared by SEER\*Prep.

**Rates and Trends** require that population data be associated with the case data. You must supply your own population data (i.e. the populations already in SEER\*Stat can not be used with the user created data). Population data for the U.S. for 1969+ are available on the SEER Program's web site under the topic U.S. Populations.

**Survival** calculations can only be performed on incidence databases. For meaningful survival there must be follow up information in the data.

SEER\*Stat's **Case Listing** functionality can be utilized with any data prepared by SEER\*Prep.



# Getting Help

SEER\*Stat provides explicit help relating to all system options. The help system defines in detail how to produce the statistics/values your analysis requires. There are four ways to access the help system from SEER\*Stat:

- Choosing Contents from the Help menu
- Clicking on the Help Contents button on the tool bar
- Pressing F1 on the keyboard for context-sensitive help
- Clicking on a screen's Help button when provided

## Help Contents

There are three tabs in Help Contents: contents, index, and find. Each tab accesses the same help file.

### Contents Tab

On the Contents Tab, there is a list of topics with book icons. If you double click the book icon next to the topic you want, the book will open and more books and pages will appear. The books are the chapter and section headings. The pages are the actual help information. When you double click a page, the associated help screen will appear.

### Index Tab

To use the Index Tab, enter the first few letters of the word you want to find. If topic titles appear, click the topic you want from the index list. Then click Display to read about your selected topic. The Index Tab displays only topics in the index. Not all topics will be listed. If you cannot find your topic in the index, you should try the Find Tab.

### Find Tab

To use the Find Tab, enter the first few letters of the word you want. From the Find Tab, the help system looks for the word in the text of the help file regardless of whether it is in the index. Click on words that apply to your topic from the first list. Then, click on a topic from the second list. Finally, click Display to read about your topic.

## Context-Sensitive Help

Context-sensitive help is the most direct way to get help for anything you see on a window. Pressing **F1** on your keyboard or clicking on a Help button will bring context-sensitive help for whatever window is active.

## Technical Support

For an up to date list of known problems, frequently asked questions (FAQ's), and resolutions, please consult the SEER\*Stat Web page at:

[www.seer.cancer.gov/ScientificSystems/SEERStat](http://www.seer.cancer.gov/ScientificSystems/SEERStat)



For technical support beyond the help system, please send email to IMS at:

[seerstat@ims.nci.nih.gov](mailto:seerstat@ims.nci.nih.gov)

# Frequency/Rate Sessions

## Open a Frequency/Rate Session

To open a Frequency/Rate Session, you may do any of the following.

- Click  to open a new Frequency/Rate Session or use the New command on the File menu.
- Click  to open a previously saved Frequency/Rate Session or use the Open command on the File menu. '.si' is the Frequency/Rate Session file extension.
- Recall the session from a saved matrix file by opening the matrix and using the View Session command on the Matrix menu.

The variables available for an opened session include the standard variables and all current user-defined variables saved to the database selected on the Data Tab. Sessions may also include session specific user-defined variables.

## Frequency/Rate Session Overview

To set up a Frequency/Rate Session, work from left to right across the tabs, and from top to bottom within each tab.

### 1 Frequency/Rate Data Tab

On the Frequency/Rate Data Tab you will select the database on which you will base your analysis. You will also select the statistics and adjust any parameters necessary, including p-values, standard million, age variable, and trend variable.

### 2 Frequency/Rate Selection Tab

The search statements on the Frequency/Rate Selection Tab define which subset of cases you wish to use in your analysis. When making selections, start in the top box and work your way down the window. Search statements are created on the Selection Window. To access the Selection Window, click the Edit button next to the text box. For further instructions on making selections see Working with Search Statements.

### 3 Frequency/Rate Table Tab

On the Frequency/Rate Table Tab you may select the row, column, and page variables for your output matrix.


#### 4 Frequency/Rate Output Tab

On the Frequency/Rate Output Tab you can:

- Add a title to your output matrix
- Display row or column percentages for frequency calculations
- Remove statistics with low case counts
- Display standard errors and confidence intervals for rates or trends
- Display p-values for EAPC significance testing
- Display rates as cases per some power of 10

When you are ready to execute, click  to create the output matrix .

## Save a Frequency/Rate Session

- 1 Click on the Frequency/Rate Session Window to activate it.
- 2 Click  or use the Save or Save As commands on the File menu.

The Frequency/Rate Session file may be saved at any time, even if it has not been executed. A saved Frequency/Rate Session can be retrieved and modified at a later time.

It is not necessary to keep both the session and its output matrix. A session may be executed again to recreate the matrix. Alternatively, since the session and matrix are saved together, the session can be recalled from the matrix.

## Frequency/Rate Data Tab

On the Frequency/Rate Data Tab you will select the database on which you will base your analysis. You will also select the statistics and adjust any parameters necessary, including p-values, standard million, age variable, and trend variable.

### Database

You may choose a database like SEER Public-Use for your analysis. Rates or trends may be calculated with a database that has a population file associated with it. If there isn't an associated population file, only frequencies and percentages can be calculated. You cannot select a population file for a database nor can you change population files. The first database listed is the default.

### Statistics

Frequencies are the default statistics selection. If the database you have chosen is associated with a population file, then you will be able to select any of the following statistics.

<b>Frequencies</b>	These are simply a count of cases. Row or column percentages may also be calculated. This is the default statistic. See the Frequency/Rate Output Tab for more information.
<b>Crude Rates</b>	These are the count of cases divided by the associated population count and multiplied by a power of ten (100,000 by default). When you select crude rates, the output matrix will display the counts and populations along with the rates.
<b>Age-adjusted Rates</b>	These are the crude rates which have been age-adjusted (or weighted by age group) to the standard million file of your choice. When you select age-adjusted rates, the output matrix will display the rates with the counts and population counts used to calculate them.
<b>Trends</b>	These are changes in rates (crude or age-adjusted) over time. They are expressed in 2 forms: the percent change (PC) and the estimated annual percent change (EAPC).
<b>Parameters</b>	The statistic you choose for your analysis determines which parameters are needed. For example, you need the standard million and age variable parameters only if you are creating age-adjusted rates. On the Frequency/Rate Data Tab, you have the opportunity to change the settings for the p-values, the standard million year, the age variable, and trend variable as needed.

## Database for a Frequency/Rate Session

You may choose a database like SEER Public-Use for your analysis. Rates or trends may be calculated with a database which has a population file associated with it. If there isn't an associated population file, only frequencies and percentages can be calculated. You cannot select a population file for a database nor can you change population files. The first database listed is the default.

### Using Your Own Data

If you would like to use your own data with SEER\*Stat, the database must be created using the SEER\*Prep software. SEER\*Prep is available on the SEER Program's web site under the topic of Scientific Systems:

[www.seer.cancer.gov](http://www.seer.cancer.gov)

## Statistic for a Frequency/Rate Session

From the list of statistics on the Frequency/Rate Data Tab, choose one for your analysis. If you have chosen a database without an associated population file, only frequencies may be calculated. The statistic you choose will affect the parameters you need to change. All statistics will be displayed in the output matrix.

### Frequencies

These are simply a count of cases. Row or column percentages may also be calculated. See the Frequency/Rate Output Tab for more information.

### Crude Rates

These are the count of cases divided by the associated population count and multiplied by a power of ten (100,000 by default). When you select crude rates, the output matrix will display the counts and populations along with the rates.

### Age-adjusted Rates

These are the crude rates which have been age-adjusted (or weighted by age group) to the standard million file of your choice. When you select age-adjusted rates, the output matrix will display the rates with the counts and population counts used to calculate them.

### Trends (crude or age-adjusted)

These are changes in rates (crude or age-adjusted) over time. They are expressed in 2 forms: the percent change (PC) and the estimated annual percent change (EAPC).

- The PC is the difference of the average rate of the first two years and the average rate of the last two years divided by the average rate of the first two years.
- The EAPC is calculated by fitting a regression line to the natural logarithm of the rates  $r$  using the calendar year  $x$  as a regressor variable, i.e.,  $y = mx + b$  where  $y = \ln r$ .

## Statistics Requiring a Population File

Rates and trends can be calculated for databases with both a case and population file attributed to them. Population files are matched to case files by the following variables:

Age Recode

Sex

Race Recode A

Year of Diagnosis (incidence) or Year of Death (mortality)

Registry or State

County

You should choose the selection and table variables carefully to ensure that the rates or trends are calculated correctly. If you select cases based on a variable not listed above, the population file will not have the same composition as the case file. This can cause rates and trends which cannot be interpreted. For example, the selection “Age of Diagnosis = 30-39” results in a case selection of only 30 to 39 year-olds, but the population file will contain all ages. The rates will be significantly lower than expected since the denominator (population) will include people that were not eligible to be cases.

## **Parameters for a Frequency/Rate Session**

The statistic you choose for your analysis determines which parameters are needed. For example, you need the standard million and age variable parameters only if you are creating age-adjusted rates. On the Frequency/Rate Data Tab, you have the opportunity to change the settings for the p-values, the standard million year, the age variable, and trend variable as needed.

### **P-Values for a Frequency/Rate Session**

On the Frequency/Rate Data Tab, the p-values are used in the calculation of confidence intervals and in significance testing. In SEER\*Stat, there are actually four different p-values defined, one for each of the following calculations:

- Confidence intervals for rates
- Confidence intervals for EAPCs
- Significance testing of EAPCs to 0
- Significance testing of a EAPC to another EAPC

You may choose to calculate confidence intervals and display p-values used for significance testing from the Frequency/Rate Output Tab.

The default value for all p-values is 0.05, unless you reset the default. This is equivalent to a 95% confidence interval or 0.05 significance. If you want to change any p-value temporarily or permanently, click the Edit Button.

### **Edit P-Values Used in a Frequency/Rate Session**

You may change the p-value which appears on the Frequency/Rate Data Tab for the current session by clicking the Edit Button. You may then edit the text boxes on the Level of Significance (P-Values) Window. Each text box represents one of the p-values.

<b>Rate</b>	Confidence intervals for rates = $1 - p\text{-value}$
<b>EAPC</b>	Confidence intervals for EAPCs = $1 - p\text{-value}$
<b>EAPC to 0</b>	Comparison is significant to the p-value
<b>EAPC to EAPC</b>	Comparison is significant to the p-value

When you are finished editing, click on OK.

### **Default**

Click this button to return to the default settings for the p-values. Initially, the default setting for each p-value is 0.05.

### **Set Default**

Once you have edited the p-values, you may set the current values as the default p-values by clicking on this button. This may be useful if you use a certain set of p-values regularly. These p-values will now be the default whenever you open the a Frequency/Rate Session.

## **Standard Million for Frequency/Rate Session**

On the Frequency/Rate Data Tab, the standard million provides the proportions used as weights to create age-adjusted rates. It contains the age distribution for a standardized population. SEER\*Stat gives you the choice of the decennial U.S. Census populations from 1940 to 1990, an estimation for U.S. 2000, Canadian 1991, or the World population. By default, the 1970 U.S. Census population is chosen. To change the standard million for your analysis, select the standard million that you want from the list box.

## **Age Variable for a Frequency/Rate Session**

On the Frequency/Rate Data Tab, the age variable defines the number of age-groups used in creating age-adjusted rates and the definitions of those age-groups.

For the SEER Public-Use database, the Age Recode variable is the default with 5-year age-groups. To choose a different arrangement, make a user-defined variable based on Age Recode. This newly made variable will appear in the Age Variable list box as long as there are no overlapping or repeated age-groups. If you exclude an age-group, it will be included as its own grouping in the age-adjusting calculations.

## **Trend Variable for a Frequency/Rate Session**

On the Frequency/Rate Data Tab, the trend variable defines the year groupings for which SEER\*Stat computes the trend. Each grouping in the trend variable must contain at least three consecutive years. For each correctly defined grouping, the trend calculation produces the percent



change (PC), the estimated annual percent change (EAPC), and compares the EAPC to 0 for significance. Additionally, the rates for the years included in these groupings are displayed in the output matrix.

SEER\*Stat can also compute the significance of the comparison between the base EAPC to other EAPCs. The first grouping listed in the dictionary's definition of the trend variable serves as the base EAPC to which all other EAPCs are compared. If a grouping has multiple years overlapping the base year grouping, this significance test cannot be made.

If you wish to change the year groupings for your analysis, create a user-defined variable based on the standard trend variable for your database. For the SEER Public-Use database, the Year of Diagnosis variable is the default trend variable.

## Frequency/Rate Selection Tab

When creating a session, the search statements on the Frequency/Rate Selection Tab define for SEER\*Stat which subset of cases you wish to use in your analysis. There are three selection boxes on the Frequency/Rate Selection Tab:

- 1 Age at Diagnosis (Std Mil, Pop, Case Files)
- 2 Race, Sex, Year Dx, Registry, County (Pop, Case Files)
- 3 Other (Case Files)

Each box contains a separate set of variables. No variable is repeated among the boxes. When making selections, start at the top and work your way down the window. Frequency only databases may have only one or two of the above boxes active. If the name of the box has changed to "Not Applicable", then no variables exist for this box and the Edit Button associated with the box is disabled.

Search statements are created on the Selection Window. To access the Selection Window click the Edit Button next to the text box. If there are no search statements on the Selection Window, then all cases in the database will be selected. Therefore, you should only make search statements for variables used to exclude cases.

To clear all search statements from a selection box, click the Clear Button.

For further instructions on making selections see Working with Search Statements and Tips for Accurate Search Statements.

## Age at Diagnosis Selection for a Frequency/Rate Session

The Age at Diagnosis (Std Mil, Pop, Case Files) box on the Frequency/Rate Selection Tab consists of the variables found in the

Standard Million, Population, and Case files. If you are using the SEER Public-Use database, the only variable found in all three files is Age Recode and any user-defined variations of Age Recode. When making selections based on age at diagnosis, you should make them here, so that the appropriate subsets may be made from the population and standard million files as well as the case file.

## **Race, Sex, Year Dx, Registry, County Selection for a Frequency/Rate Session**

The Race, Sex, Year Dx, Registry, County (Pop, Case Files) box on the Frequency/Rate Selection Tab consists of all variables found in the population and case files, but not in the Standard Million file. If you are using the SEER Public-Use database, these variables are County, Race Recode A, SEER Registry, Registry/County, Sex and Year of Diagnosis. All selections by race, sex, year of diagnosis, registry, or county should be made in this box, so the appropriate subset may be made from the population file as well as the case file.

## **Other Case Variables Selection for a Frequency/Rate Session**

The Other (Case Files) box on the Frequency/Rate Selection Tab consists of all variables that are found only in the case file. If you are using the SEER Public-Use database, notice that there may be variables for race, age, and year of diagnosis available in the Other selection box. Be very careful choosing these variables in your case selection if you are calculating any kind of rate. The population and standard million files will not reflect your case subset if you make the age, race, or year of diagnosis selections from the Other (Case Files) box.

## **Frequency/Rate Table Tab**

On the Frequency/Rate Table Tab you may select the row, column, and page variables for your output matrix. This can be a powerful tool for creating exactly the output you are seeking. For example, using the SEER Public-Use database, if you want to compare the yearly diagnosis rates by race, select Year of Diagnosis as a row variable and Race Recode A as a column variable. If you want the rates by both race and sex, add Sex as the second column variable or create a merged variable. Select SEER Registry as a page variable to compare the rates within each registry. Each addition to the table variables creates more detail in the output matrix. Be careful when choosing table variables to avoid inadvertent exclusions.

### **Page, Column, and Row Lists**

These are lists of the variables for the pages, columns, and rows of your output matrix.

### **Remove Buttons**

Click **×** to remove a table variable.

### **Up and Dn Buttons**

Click **↑** or **↓** to change the order of the table variables. The order of variables in the output matrix is determined by the order of variables in this window.

### **Variables**

The variables in the database are listed in categories. Each category is named for the variables found in it. To expand a category, double click the category label or click the plus sign (+) next to it. This will reveal the variables contained in the category. Single-click variable names to select them. To view the definition of a variable in the dictionary, double-click the variable name and click the create or modify buttons on the Dictionary Window.

### **Row, Column, and Page Buttons**

Click these buttons to add the selected variable to the row, column, or page list.

### **Find Button**

If you are looking for a particular variable, but cannot find it within the variable folders, click Find to bring up the Find Window.

## **Add a Table Variable to a Frequency/Rate Session**

- 1 Click on a variable in the Variable section of Frequency/Rate Table Tab or click the Find Button to search for the variable you want.
- 2 Click Row, Column, or Page. The name of your variable will appear in the appropriate list. You can specify up to eight page variables, five row variables, and five column variables. Realize that no more than 256,000 cells per page and up to 32,767 pages may be created for a single output matrix.

### **Tip**

Check the Dictionary for the values before adding the variable to your table definition. Sometimes the subset of cases created on the Frequency/Rate Selection Tab will make variable groupings meaningless or wrong. Instead create a user-defined or merged variable which reflects the subset of cases you wish to display.

## Order the Table Variables in a Frequency/Rate Session

The order of variables in the output matrix is determined by the order of variables in each list on the Table Tab. To change the order of variables:

- 1 From the Frequency/Rate Table Tab, click the variable you want to move within the list (Row, Column, or Page).
- 2 Use ↑ and ↓ , the Move Up and Move Down commands from the Edit menu, or Ctrl+Up Arrow and Ctrl+Down Arrow on the keyboard to move the variable up or down within the list.

Variables may be moved only within a single list. To move a variable from one list to another (for example, Row to Column) you must remove the variable from one list and add it to the other.

### Note

Variables in a single list may also be reordered after the output matrix has been created by following these steps:

- 1 Select a row or column of the Frequency/Rate output matrix.
- 2 Click on right mouse button.
- 3 Select Order from the menu which appears. The Order Labels Window will appear.

## Remove a Table Variable from a Frequency/Rate Session

- 1 From the Frequency/Rate Table Tab, click the variable you want to remove from the page, row, or column list.
- 2 Click the × button, select Delete/Remove from the Edit menu, or press the Delete key on your keyboard.

## Frequency/Rate Output Tab

On the Frequency/Rate Output Tab you can:

### Add a title to your output matrix

If you wish to add a title to your output matrix, simply type your title in the text box provided. When the matrix is printed the text will be centered on the page, but will not wrap to a second line without carriage returns. To get a multi-lined title, press Enter at the end of each line of text.

### Display row or column percentages for frequency calculations

When you are calculating frequencies, you may also select to display row or column percentages. The percentages represent the distribution of

cases within the row or column. The percentages and row or column totals are displayed on the output matrix.

#### **Remove statistics with low case counts**

Selecting this option will suppress all statistics where the case count is less than the specified minimum. The cells will instead have a footnote explaining why they are empty.

#### **Display standard errors and confidence intervals for rates or trends**

When this option is selected, the standard errors and confidence intervals for the rates will be calculated and displayed in the output matrix. If trends are being calculated, confidence intervals for the EAPC will also be included.

#### **Display p-values for EAPC significance testing**


Selecting this option allows the p-values used for EAPC significance testing to be displayed in the output matrix. This option is only available when trends are being calculated. A footnote in the output matrix explains that the p-values are listed in the Standard Error column of the EAPC rows.

#### **Display rates as cases per some power of 10**

From the list box choose the power of 10 by which you want the crude rates multiplied.

## **Title for a Frequency/Rate Matrix**

If you wish to add a title to your output matrix, simply type your title in the text box provided on the Frequency/Rate Output Tab. When the matrix is printed the text will be centered on the page, but will not wrap to a second line without carriage returns. To create a multi-lined title, press Enter at the end of each line of text.

Once the matrix has been made, you can see the entire title by clicking on  or choosing Options from the Matrix menu. The title may be added or edited on the Frequency/Rate Matrix Options Window. Only the first line of the title will be displayed on the SEER\*Stat matrix window. To see how the matrix title will print, select Print Preview from the File menu.

## **Percentages for Frequency/Rate Sessions**

When you are calculating frequencies, you may also select to display row or column percentages on the Frequency/Rate Output Tab. The percentages represent the distribution of cases within the row or column. In particular, row percentages assume that the row of a single page adds up to 100% and the column percentages assume that the column of a single page adds to 100%. The percentages and row or column totals are

displayed on the output matrix. There is no facility to calculate percentages across pages of the output matrix.

When table variables have overlapping groupings, an error message will appear to warn you that the percentages will not make sense. If you do not think that you should have gotten the error message, check the definition of the groupings for the table variable in the database dictionary. If any value of the variable is found in more than one grouping, create a user-defined variable which does not have overlapping groupings.

For example, if you chose Sex as a column variable and selected to display row percentages, you would receive an error because the group “Male and Female” and the group “Male” both include the value “Male”. To correct this, you would follow these steps:

- 1 Open the dictionary.
- 2 Select Sex and click the Create button.
- 3 Rename the Sex variable to Sex (m,f).
- 4 Delete the grouping “Male and Female”.
- 5 Save the user-defined variable Sex (m,f).
- 6 On the Table Tab, remove Sex from the row variables list.
- 7 Add the user-defined variable Sex (m,f) to the row variables list.
- 8 Execute the session; no error appears.

## Hide Statistics When Less Than \_\_\_\_ Cases

Selecting this option from the Frequency/Rate Output Tab will suppress all statistics where the case count is less than the specified minimum. The cells will instead have a footnote explaining why they are empty. This feature was included in SEER\*Stat to encourage the protection of patient privacy.

## Display Standard Errors and Confidence Intervals

Selecting this option from the **Frequency/Rate Output Tab** causes SEER\*Stat to calculate the standard errors and confidence intervals for the rates. If trends are being calculated, confidence intervals for the EAPC will also be included. The degree of confidence is determined by the p-value. It is equal to one minus the p-value. For example, the p-value equal to 0.05 yields 95% confidence intervals. The p-values may be edited on the Frequency/Rate Data Tab.

### Definitions

The standard error of a rate is a measure of the sampling variability of the rate.

A confidence interval is a range of values that has a specified probability of containing the rate or trend. The 95% (p-value = .05) and 99% (p-value = .01) confidence intervals are the most commonly used.

## **Display p-Values for EAPC Significance Testing**

Selecting this option from the Frequency/Rate Output Tab allows the p-values used for EAPC significance testing to be displayed in the output matrix. This option is only available when trends are being calculated. A footnote in the output matrix explains that the p-values are listed in the Standard Error column of the EAPC rows.

## **Display Rates as Cases Per \_\_\_\_\_**

SEER\*Stat displays rates as cases per some power of ten population. From the list box on the Frequency/Rate Output Tab choose the power of 10 by which you want the rates multiplied. The default value may be changed on the Preferences Window.

You can select:

1,000

10,000

100,000

1,000,000







# Survival Sessions

## Open a Survival Session

To open a survival session, you may do any of the following.

- Click  to open a new survival session or use the New command from the File menu.
- Click  to select a previously saved survival session or use the Open command from the File menu. '.ss' is the Survival Session file extension.
- Recall the session from a saved matrix file by opening the matrix and using the View Session command on the Matrix menu.

The variables available to an opened session include the standard variables and all current user-defined variables saved to the database selected on the Data Tab. Sessions may also include session specific user-defined variables.

## Survival Session Overview

To set up a survival analysis, work from top to bottom through each tab in order from left to right.

### 1 Survival Data Tab

On the Survival Data Tab you will select the database and statistics on which you will base your analysis. If you choose to calculate relative survival rates, you will also select the expected rates table.

### 2 Survival Selection Tab

From the Survival Selection Tab, you may define which subset of cases you wish to use in your analysis. There are check boxes for standard selections made for survival runs on the tab. In addition, a search statement may be created to select the cohort of cases you will analyze. For further instructions on making selections see Working with Search Statements.

### 3 Survival Table Tab

On the Survival Table Tab you will select the page variables for your survival analysis. Page variables dictate how the cases selected for the analysis are to be grouped into life tables. Each life table is defined by a specific combination of page variable values called the table definition. Be careful when making page variables to avoid inadvertent exclusions.


#### 4 Survival Parameters Tab

Use the Survival Parameters Tab to specify the dates, intervals, and vital status involved with the calculation of survival time. These parameters establish how case survival time will be defined and the intervals that will be used in the analysis.


#### 5 Survival Output Tab

On the Survival Output Tab you can:

- Add a title to your output matrix
- Display Tables: Standard Life and Cumulative Summary
- Flag high relative cumulative standard errors
- Remove pages with low case counts
- Adjust relative survival rates over 1.0
- Adjust increasing relative survival rates
- Adjust for heterogeneity in withdrawal

When the Survival Session is complete, click  to create the output matrix.

## Save a Survival Session

- 1 Click on the survival session to activate it.
- 2 Click  or use the Save or Save As commands on the File menu.

The survival session file may be saved at any time, even if it was not executed. A saved survival session file can be retrieved and modified at a later time.

It is not necessary to keep both the session and its output matrix. Sessions may be executed again to recreate the matrix. It is also possible to recall the session from the saved matrix file, for the session is saved with the matrix. User-defined and merged variables which are not saved to the dictionary and are not used in the saved version of a session or matrix are lost.

## Survival Data Tab

On the Survival Data Tab you will select the database and statistics on which you will base your survival analysis. If you choose to calculate relative survival rates, you may also select the expected rates tables.

## Database

Click the incidence database you wish to use for your survival analysis. Only one input data file may be selected at a time. Choose the incidence database for your analysis at the beginning of your survival session. Each incidence database has a unique data dictionary which lists the available variables and the values. These variables are used throughout the survival session.

## Statistics

When choosing the statistics you want, keep in mind the type of rates you need. You may select either:

<b>Observed and Relative Survival</b>	Calculates the observed, expected, and relative survival rates.
<b>Observed Survival Only</b>	Calculates the observed survival rates alone.

## Expected Rates Tables

An expected rates table is needed only when relative survival statistics have been chosen. Expected rates are obtained from the expected rate table by matching the cohort cases by race, sex, age, and date at which the age was coded. Together the observed and expected rates are used to generate relative survival rates.

# Database for a Survival Session

From the Survival Data Tab, click the incidence database you wish to use for your survival analysis. Only one input data file may be selected at a time. Choose the incidence database for your analysis at the beginning of your survival session. Each incidence database has a unique data dictionary which lists the available variables and the values. These variables are used throughout the survival session.

## Using Your Own Data

If you would like to use your own data with SEER\*Stat, it must be processed using SEER\*Prep. For meaningful survival statistics, you must have follow up information included in your incidence data. SEER\*Prep is available on the SEER Program's web site under the topic of Scientific Systems:

[www.seer.cancer.gov](http://www.seer.cancer.gov)

# Survival Rates Definitions

When choosing the statistics you want from the Survival Data Tab, keep in mind the type of rates you need. Here are the definitions of the three rate types available through a SEER\*Stat Survival Session. To calculate all three types of rates, you should choose Observed and Relative Survival.

### **Observed Survival Rates**

Observed survival is the actual probability of surviving a specified time interval as calculated from the cohort of cancer cases. You can choose to calculate the observed survival rates alone by choosing Observed Survival Only.

### **Expected Survival Rates**

Expected survival is the probability of surviving the specified time interval in a general population. Expected probabilities (rates) generated from the U.S. population have been included with SEER\*Stat. They are matched to the cohort cases by race, sex, age, and date at which age was coded. You may also prepare your own expected rates using the SEER\*Prep program.

### **Relative Survival Rates**

Relative survival is the observed survival probability for the specified time interval adjusted for the expected survival. The relative survival rate adjustment accounts for the general survival rate of the U.S. population for that race, sex, age and date at which the age was coded.

## **Observed and Relative Survival**

By selecting Observed and Relative Survival from the Survival Data Tab, the following statistics will be shown for each survival interval in the output life tables.

- Number of cases entering the interval alive
- Number of cases dying, lost to follow up, or withdrawn alive\* during the interval
- The adjusted number of alive\* cases for each interval
- Observed survival, expected survival, and relative survival rates for the interval
- Cumulative observed, expected, and relative survival rates
- Standard errors for the observed survival, cumulative observed survival, relative survival, and cumulative relative survival

\* These two statistics do not appear in the output matrix unless selected from the Options Window from the Matrix menu.

## **Observed Survival Only**

By selecting Observed Survival Only from the Survival Data Tab, the following statistics will be shown for each survival interval in the output life tables.

- Number of cases entering the interval alive

- Number of cases dying, lost to follow up, or withdrawn alive\* during the interval.
  - The adjusted number of alive\* cases for each interval
  - Observed survival for the interval and the cumulative observed survival
  - Standard errors for the observed survival and cumulative survival
- \* These two statistics do not appear in the output matrix unless selected from the Options Window from the Matrix menu.

## Expected Rate Table

Choose the expected rate table from the Survival Data Tab for your analysis. The expected rate table is needed only when relative survival statistics have been chosen. Expected rates are obtained from the expected rate table by matching the cohort cases by race, sex, age, and date at which the age was coded. Together the observed and expected rates are used to generate relative survival rates.

The expected rate tables that come with SEER\*Stat contain the expected probabilities of survival generated from the general U.S. population. For the SEER Public-Use databases, the default expected rate table was used to create survival statistics in the SEER Cancer Statistics Review. Specific information about the expected rate tables delivered with the SEER\*Stat CDs follows.

**U.S. 1970, 1980, 1990 (White, Black, Other)** matches to the cohort cases using the SEER Race Recode A variable and applies the 1970 rates to 1973-1975 ages, the 1980 rates to the 1976-1985 ages, and 1990 rates to the ages recorded in 1986 through the year of latest diagnosis.

**U.S. 1970, 1980 (Expanded Races)** matches to the cohort cases using the SEER Race Recode B variable and applies 1970 rates to the 1973-1975 ages and the 1980 rates to the ages recorded in 1976 and after.

It is also possible to prepare your own expected rate tables using SEER\*Prep. SEER\*Prep is available on the SEER Program's web site under the topic of Scientific Systems:

[www.seer.cancer.gov](http://www.seer.cancer.gov)

To view the expected rate tables, follow these steps:

- 1 Open a Frequency/Rate Session.
- 2 On the Frequency/Rate Session Data Tab, select the expected rates database you are interested in viewing.
- 3 On the Frequency/Rate Session Table Tab, set the row variable to be Age and the column variables to be Year, Sex, and Race.
- 4 Execute the Frequency/Rate Session.

**Note**

You may make selections if you wish on the Frequency/Rate Selection Tab. Nevertheless, it is very important to select all variables of the expected rates database as table variables. If you do not, then the rates will be summed, which is meaningless.

## Survival Selection Tab

When creating a survival session, you may define the subset of cases you wish to use in your analysis. There are two regions on the Survival Selection Tab from top to bottom:

- 1 Standard Selections which includes selections made due to the expected rates table that you are using

The Standard Selections box contains a set of case selection criteria commonly associated with survival analyses. These statements select or exclude case records based on the criteria stated. Each standard selection or exclusion can be accepted or rejected by clicking on it.

- 2 Case Selection

Typically, you will confine your analysis to a subset of cases with a certain cancer or particular demographics. When making selections, start at the top and work your way down the window. For the case selection, the search statement is created on the Selection Window. To access the Selection Window, click the Edit Button next to the box. If there are no search statements on the Selection Window, then all cases in the database will be selected. Therefore, you should only make search statements for variables used to exclude cases.

## Standard Survival Selections

The Standard Selections box on the Survival Selection Tab contains a set of case selection criteria commonly associated with survival analyses. These statements select or exclude case records based on the criteria stated. Each standard selection or exclusion can be accepted or rejected by clicking on it.

When a new survival session is started, all but one of the standard selections or exclusions will be automatically selected. These default selections represent the standard selections most commonly used for a survival analysis. By marking a standard selection, all cases which fall into the category will be selected or excluded as advertised.

### Select Only Microscopically Confirmed

Selecting this option from the Survival Selection Tab will include all cases which were microscopically confirmed. The Diagnostic Confirmation variable indicates whether there was microscopic confirmation of the

morphology of the cancer. The case will be included in the analysis if the Diagnostic Confirmation variable indicates that there was microscopic confirmation of the morphology.

### **Select Only Actively Followed**

Selecting this option from the Survival Selection Tab will include only cases which were actively pursued for follow up information. The Type of Follow Up variable indicates whether the case was actively followed. If the Type of Follow Up variable indicates that the case was actively followed, the case will be included in the analysis.

### **Select Only Malignant Behavior**

Selecting this option from the Survival Selection Tab will include only cases with a malignant behavior. The Behavior variable is part of the morphology of cancer. The SEER program only collects information on in situ and malignant cancers. If the behavior code is malignant, the case will be included in the analysis.

### **Exclude All Second and Later Primaries**

Selecting this option from the Survival Selection Tab will exclude all cases which are second or later primaries. The Sequence Number variable indicates the chronology of diagnoses of all reportable primary malignant and/or in situ cancers over the known history of the patient. The case will be excluded from the analysis if it is not the first primary cancer.

### **Exclude All Death Certificate and Autopsy**

Selecting this option from the Survival Selection Tab will exclude all cases whose cancer was reported through only a death certificate or autopsy. The Type of Reporting Source variable indicates where the information for the cancer was obtained. If the Type Of Reporting Source variable indicates that a death certificate or autopsy was the only source of information for the case, the case will be excluded from the analysis.

## **Expected Rate Table Standard Survival Selections**

Some standard exclusions on the Survival Selection Tab are selected when relative survival is calculated so that the cases will match the expected rate tables. When you calculate observed survival only, you may choose not to make these exclusions. Making these exclusions will allow you to produce the same results for observed survival and relative survival sessions.

### **Exclude All Unknown Races**

Selecting this option from the Survival Selection Tab will exclude all cases with a missing or unknown race. This exclusion is always selected when

Relative Survival is being calculated because there are no rates for cases with unknown race in the expected rate table you are using. Since expected rates are matched to incidence cases by race, age, sex, and date at which the age was coded, those cases with unknown or invalid information in any of these variables cannot be matched and must be excluded.

### **Exclude All Unknown Ages**

Selecting this option from the Survival Selection Tab will exclude all cases with ages outside the boundaries of the expected rate table or with unknown ages. This exclusion is always selected when Relative Survival is being calculated because there are no rates for cases with unknown age in the expected rate table you are using. Since expected rates are matched to incidence cases by race, age, sex, and date at which the age was coded, those cases with unknown or invalid information in any of these variables cannot be matched and must be excluded.

### **Exclude All Sex not Male or Female**

Selecting this option from the Survival Selection Tab will exclude all cases with values for sex other than 'male' or 'female.' This exclusion is always selected when Relative Survival is being calculated because there are no rates for cases with sex other than male or female in the expected rate table you are using. Since expected rates are matched to incidence cases by race, age, sex, and date at which the age was coded, those cases with unknown or invalid information in any of these variables cannot be matched and must be excluded.

### **Exclude All Invalid Age Dates**

Selecting this option from the Survival Selection Tab will exclude all cases with an age date, the date at which the age was coded, outside the boundaries of the expected rate table or with an unknown age date. In SEER, the age date is the date of diagnosis. This exclusion is always selected when Relative Survival is being calculated because there are no rates for unknown age dates in the expected rate table you are using. Since expected rates are matched to incidence cases by race, age, sex, and the age date, those cases with unknown or invalid information in any of these variables cannot be matched and must be excluded.

## **Case Selection for a Survival Session**

You will usually want to confine your analysis to a subset of cases with a certain cancer or particular demographics. To select the case subset for your analysis, click the Edit button on the Survival Selection Tab. This will bring up the Selection Window. Remember that if you have no search statements, you will have selected all the cases in the database which



conform to the standard selections you have chosen. Therefore, make search statements for variables used to exclude cases.

## Survival Table Tab

On the Survival Table Tab you will select the page variables for your survival analysis. Page variables dictate how the cases selected for the analysis are to be grouped into life tables. Each life table is defined by a specific combination of page variable values called the table definition. Be careful when making page variables to avoid inadvertent exclusions.

### Page

This is a list of the page variables for your survival statistics.

### Variables

The variables in the database are listed in categories. Each category is named for the variables found in it. To expand a category, double click the category label or click the plus sign (+) next to it. This will reveal the variables contained in the category. Single click a variable name to select it. To view the definition of a variable in the dictionary, double-click the variable name and click the Create or Modify buttons on the Dictionary Window.

### Find Button

If you are looking for a particular variable, but cannot find it within the variable folders, click Find to bring up the Find Window .

### Remove Button

Click here to remove a page variable.

### Move Up and Move Down Button

Click here to change the order the page variables. The order of page variables in the output matrix is determined by the order of variables in this window.

### Add Button

Click here to add a page variable.

### Tip

Consider creating a merged variable for all the variable value combinations you are interested in displaying. This way you can display all of your combinations in the same the Cumulative Summary Table.

## Add a Page Variable to a Survival Session

- 1 Click on a variable from the Variable section of the Survival Table Tab or click the Find button to select the variable you want.

- 2 Click the Add button. The name of your variable will be listed as a page variable. You can specify up to eight page variables.

### **Tips**

Check the Dictionary for the variable groupings before adding the variable to your table definition. Sometimes the subset of cases created on the Survival Selection Tab will make the variable groupings meaningless or wrong. Instead, create a user-defined variable which reflects the subset cases.

## **Order the Page Variables in a Survival Session**

The order of page variables in the output matrix are determined by the order of variables on the Survival Table Tab. To change the order of variables:

- 1 From the Survival Table Tab, click the page variable you want to move.
- 2 Click the Move Up or Move Down buttons, the commands from the Edit menu, or use Cntrl+Up and Cntrl+Down on the keyboard.

## **Remove a Page Variable from a Survival Session**

- 1 From the Survival Table Tab, click the page variable you want to remove.
- 2 Click the Remove button, select Delete/Remove from the Edit menu, or press the Delete key.

## **Avoid Inadvertent Exclusions**

A case is excluded from your analysis if it does not pass the selection criteria of the Selection Tab and match at least one table definition of the Table Tab. For best results, you should select your case subset while you are on the Selection Tab rather than relying on the table definitions to exclude cases. After the output matrix is created, SEER\*Stat displays a count of cases excluded because they do not satisfy any table definitions in the Information section of the Matrix Properties.

## **Survival Parameters Tab**

Use the SEER\*Stat Survival Parameters Tab to specify the dates, intervals, and vital status involved with the calculation of survival time. These parameters establish how case survival time will be defined and the intervals that will be used in the analysis.

The survival time for each case is calculated and then partitioned into the defined survival duration intervals. For every life table, SEER\*Stat calculates the number of cases alive at the beginning of each survival

interval. A case is counted as entering the first interval alive and, either survives in to the next interval, or its survival time terminates somewhere within the current interval. If the case does not survive into the next interval, it is counted as either dying during the current interval or its vital status indicates that its alive, but does not have enough survival time to go into the next interval. In the latter case, it is counted as lost to follow-up or withdrawn during the interval, depending on the relationship between the End Date of Calculation and the Lost to Follow-up Date.

Cases which are alive and have identical begin and end dates of calculation are deleted from the analysis since the survival time is zero. The number of cases for which survival could not be calculated is included in the Information section of the Matrix Properties.

## **Dates for a Survival Session**

The system calculates the case survival time from a begin date and end date stored on each record. The begin date is typically selected to be the date of diagnosis and the end date, the date of death or last contact. With each ending date of calculation, there is an associated vital status field. The vital status indicates whether the case was alive or dead at the end date.

When creating a life table, SEER\*Stat calculates the number of cases alive at the beginning of each survival interval. Each case is counted as entering the first interval as alive. The case either survives into the next interval or its survival time terminates somewhere within the current interval. If the case does not survive into the next interval, it is counted as either dying during the current interval or the vital status indicates that it is alive, but does not have enough survival time to go in the next interval. In the latter case, the case is counted as lost to follow-up or withdrawn during the interval depending on the relationship between its End Date of Calculation and the Lost to Follow-up Date.

There are two additional date parameters established on the Survival Parameter Tab which can affect the survival calculation: study cutoff date and lost to follow-up date.

### **Begin Date**

The Begin Date list box on the Survival Parameters Tab contains the date variables available to initiate the calculation of survival time for an individual case. Typically, the begin date is the date of diagnosis. If the incidence database has only one date variable suitable for the Begin Date, the list box will be fixed on the only choice. Cases with an invalid Begin Date will not be included in the analysis.

## **End Date**

The End Date list box on the Survival Parameters Tab contains the date variables available to end the calculation of survival time for an individual case. Typically, the end date is the date of last follow up or death. If the incidence database has only one date variable suitable for the End Date, the list box will be fixed on the only choice. Cases with an invalid End Date will not be included in the analysis.

## **Study Cutoff Date**

The study cutoff date specifies the date the study ended.

- 1 From the Month list box on the Survival Parameters Tab, next to the words Study Cutoff Date, click the month you want the analysis to end.
- 2 From the Year list box, next to the words Study Cutoff Date, click the year you want the analysis to end.

### **Tips**

A patient's survival time is not calculated beyond the study cutoff date. When a patient survives beyond the study cutoff date, survival time is calculated through the study cutoff date only. Sometimes there are patients which survive beyond the study cutoff date but are dead at the date of last contact. The vital status for these patients is changed to alive and their end date is set back to the study cut off date because they were alive through the end of the study.

## **Lost to Follow Up Date**

The lost to follow up date is used to determine whether cases which are alive at last contact should be considered withdrawn alive or untraced.

- 1 From the Month list box on the Survival Parameters Tab, next to the words Lost to Follow Up Date, click the month you want cases to be considered lost to follow up.
- 2 From the Year list box, next to the words Lost to Follow Up Date, click the year you want cases to be considered lost to follow up.

### **Tips**

A case which has a vital status of alive at last contact and an End Date after the Lost To Follow Up Date is considered withdrawn alive at the End Date.

A case which has a vital status of alive at last contact and an End Date before the Lost To Follow Up Date is considered untraced (lost to follow up) at the End Date. The case is untraced because there is no data available to evaluate the status during the period between the End Date and the Lost To Follow Up Date.

## Intervals for a Survival Session

A survival follow-up interval is a constant time period used to examine survival duration for cases included in a survival analysis. Every survival analysis has a constant number and length of survival intervals. The number of observation intervals and the months per interval are universal parameters which may be defined on the Survival Parameters Tab. The system provides the ability to create groupings of regular intervals called special intervals. These special intervals are used to obtain conditional survival rates. A z-test interval may also be defined.

### Number of Intervals

The Number of Intervals box on the Survival Parameters Tab indicates the maximum number of survival intervals to be displayed in the output rate tables. The default is 10 intervals. To change the number of intervals in your analysis, you may do either of the following.

- Type in the number of intervals in the box you want.
- Use the increment/decrement buttons to the right of the Number of Intervals box to increase or decrease the number of intervals.

### Months Per Interval

The Months Per Interval box on the Survival Parameters Tab indicates the length of a survival interval in months. The default is 12 months (annual interval length). To change the length of the intervals in your analysis, you may do either of the following.

- Type in the number of months you want.
- Use the increment/decrement buttons to the right of the Months Per Interval box to increase or decrease the number of months per interval.

### Special Intervals

On the Survival Parameters Tab, you may define Special Survival Follow Up Intervals which are sets of the regular intervals or specifications for conditional survival rates. A table of 10 yearly intervals might have the special interval of years 6-10, which is rate conditional on already surviving 5 years.

- 1 Determine which intervals you wish to group into a special interval.

For example, if the interval length is 12 months, to specify five year intervals which are conditional on the survival of the first five years and the first ten years, you need the intervals from 6 to 10 and 11 to 15.

- 2 Type each interval group in the Special Intervals field.

In our example, you would type the following: 6-10,11-15

## Z-Test Interval

- 1 Check that a page variable was chosen on the Survival Table Tab.  
For example if you were wanting to compare the relative survival of males to females diagnosed with lung cancer, you would want to select Sex as the page variable on the Survival Table Tab.
- 2 Click the interval number on which you want the Z-Test to end on the Survival Parameters Tab.  
For example, if the interval length is 12 months, to compare five year relative survival, you would choose interval 5.

### Tips

A Z-test is a statistical procedure for comparing the relative survival of two groups of cases. The Z-test used by SEER\*Stat compares the survival curves of two groups of cases up to a selected survival duration point (Z-test interval). Both curves must contain survival information up to (and including) the survival duration point selected.

For a detailed description of the Z-test methodology, please reference: Brown, C.C. (1983) The Statistical Comparison of Relative Survival Rates. *Biometrics* 39:941-948.

## Vital Status for a Survival Session

The Vital Status list box on the Survival Parameters Tab contains a list of all valid vital status fields associated with the selected end date of calculation. Vital Status fields indicate whether the person was alive or dead at the coded end date of calculation. Specific Vital Status fields may break down the alive or dead status into more detailed sets.

If the incidence database has only one variable which records the vital status, the list box will be fixed on the only choice. The indicator of vital status at last contact must be coded as alive or dead. Survival time will not be calculated for cases with an unknown survival status.

## Survival Output Tab

On the Survival Output Tab you can:

### Add a title to your output matrix

Type the title in the text box provided. When the matrix is printed the text will be centered on the page, but will not wrap to a second line without carriage returns. To get a multi-lined title, press Enter at the end of each line of text.

**Display Tables: Standard Life and Cumulative Summary**

You may chose to display the Standard Life Tables and/or the Cumulative Summary Table. Click the check box next to the table(s) you want. If you select to display the Cumulative Summary, do the following:

- 1 In the interval text box, enter the intervals in order from lowest to highest for which you want summary statistics.
- 2 Click the check boxes for the statistics you want summarized.

**Flag high relative cumulative standard errors**

Selecting this option will flag all cumulative relative survival rate standard errors greater than the specified percent.

**Remove pages with low case counts**

Selecting this option will suppress the creation of all survival life pages which have the number of cases entering the page less than the specified minimum.

**Adjust relative survival rates over 1.0**

When this box is checked, any calculated relative rate which exceeds 1.00 will be adjusted down to 1.00 on the output tables.

**Adjust increasing relative survival rates**


When this box is checked, any cumulative relative survival rate which exceeds the rate in the previous interval will be adjusted down to equal the previous rate.

**Adjust for heterogeneity in withdrawal**

Adjust the calculation of expected survival to account for heterogeneity in patient withdrawal due to end of follow-up.

## Title for a Survival Matrix

If you wish to add a title to your output matrix, simply type your title in the text box provided on the Survival Output Tab. When the matrix is printed the text will be centered on the page, but will not wrap to a second line without carriage returns. To get a multi-lined title, press Enter at the end of each line of text.

Once the matrix has been made, you can see the entire title by clicking on  or choosing Options from the Matrix menu. Only the first line of the title will be displayed on the SEER\*Stat matrix window. To see how the matrix title will print, select Print Preview from the File menu.

## Display Tables

You may chose to display the Standard Life Tables and/or the Cumulative Summary.

- 1 To choose a table, click the check box next to the table you want.
- 2 If you select to display the Cumulative Summary, you may enter the summary intervals you want in the interval text box and which statistics you want summarized. For example, if you wanted summary observed and relative rates at the end of 5<sup>th</sup> and 10<sup>th</sup> intervals, you would enter 5,10 in the interval text box and check the Observed check box and the Relative check box.

## Standard Life Table

When chosen from the Survival Output Tab, the Standard Life Tables record the basic survival statistics. A list of the statistics available for display on the Standard Life Table can be found under the title Observed and Relative Survival.

## Cumulative Summary

When you choose to display the Cumulative Summary on the Survival Output Tab, the selected cumulative rates and/or standard errors at each of the designated summary intervals will be displayed for each life table. Combinations of the page variables for the life tables which were chosen on the Survival Table Tab define the page and row headings. If you want a single page cumulative summary table, create a merged variable for the combinations of page variables and use it as the only page variable on the Survival Table Tab and select only one interval or only one type of rate to be summarized. When selected for display, the Cumulative Summary Tables are always at the beginning of the output matrix.

For example, suppose you define an observed and relative survival run with 10 intervals of 12 months. You then select race and sex as the page variables. Finally you specify 5,7,8 in the Cumulative Summary Interval text box and check the box next to Relative. These choices create a single summary rate page in the final matrix whose rows would correspond to the sex and race combinations which define the life tables and the columns would list the number of cases alive in the table (N) and the associated 5-year, 7-year, and 8-year relative cumulative survival rates. The intervals must be listed in order from lowest to highest.

## Flag Cum Relative SE's Greater Than \_\_\_\_ %

Selecting this option from the Survival Output Tab will flag all cumulative relative survival rate standard errors greater than the specified percent. This flag affects all cumulative summary and survival life pages. Since all standard errors are multiplied by a factor of two prior to display in the survival life tables, the standard error displayed must be divided by two and changed to a percent (multiplied by 100) in order to be compared to the flag value.



## **Suppress Pages With Less Than \_\_ Cases**

Selecting this option from the Survival Output Tab will suppress the creation of all survival life pages which have the number of cases entering the page less than the specified minimum. The affected survival life pages will appear in the survival matrix but will be empty. The empty pages are left in the final matrix in order to properly document which pages were suppressed.

## **Adjust Relative Rates over 1.0**

When calculating relative survival, the rates can be calculated properly and yet be greater than 1.00. This occurs when the actual observed survival for the case cohort has a higher survival rate than the expected rates for that same age, race, sex and date at which age was coded.

When this box is checked on the Survival Output Tab, any calculated relative rate which exceeds 1.00 will be adjusted down to 1.00 on the output tables.

## **Adjust Increasing Relative Rates**

Sometimes the cumulative relative survival rates can be calculated properly and yet be increasing over time, making it appear as if people are rising from the dead. This occurs when the actual observed survival for the case cohort decreases more slowly than the expected rates for that same age, race, sex and year group.

When this box is checked on the Survival Output Tab, any cumulative relative survival rate which exceeds the rate in the previous interval will be adjusted down to equal the previous rate.

## **Adjust for Heterogeneity in Withdrawal**



Unadjusted expected survival rates do not account for the possibility that the potential follow up times of patients are of unequal length. Unequal potential follow up time affects the estimate of expected heterogeneity between patients diagnosed with different calendar years (i.e. if the distribution of race, sex, or age change over time). The exact and simplified adjustment in the calculation of expected survival is described in detail by Hakulinen and Abeywickrama [Hakulinen T and Abeywickrama KH. A computer program package for relative survival analysis. Computer Programs in Biomedicine 1985;19:197-207].



# Case Listing Sessions

## Open a Case Listing Session

To open a Case Listing Session, you may do any of the following.

- Click  to open a new Case Listing Session or use the New command on the File menu.
- Click  to open a previously saved Case Listing Session or use the Open command on the File menu. '.sl' is the Case Listing Session file extension.
- Extract the session from a saved matrix file by opening the matrix and using the View Session command on the Matrix menu.

The variables available for a session will include the standard variables and all current user-defined variables saved to the database selected on the Data Tab. Sessions may also include session specific user-defined variables.

## Case Listing Session Overview

To set up a Case Listing Session, work from left to right across the tabs.

### 1 Case Listing Data Tab

On the Case Listing Data Tab you will select the database containing the records you wish to view.

### 2 Case Listing Selection Tab

The search statements on the Case Listing Selection Tab define the subset of cases you wish to view. Search statements are created on the Selection Window. To access the Case Selection Window, click Edit next to the text box. For further instructions on making selections see Working with Search Statements.

### 3 Case Listing Table Tab


On the Case Listing Table Tab you may select column and/or sort variables for your output matrix.

### 4 Case Listing Output Tab

On the Case Listing Output Tab you can enter a title for your output matrix.

When you are ready to execute, click  to create the output matrix .

## Save a Case Listing Session

- 1 Click on the Case Listing Session Window to activate it.
- 2 Click  or use the Save or Save As commands on the File menu.

The Case Listing Session file may be saved at any time, even if it has not been executed. A saved Case Listing Session can be retrieved and modified at a later time.

It is not necessary to keep both the session and its output matrix. A session may be executed again to recreate the matrix. Alternatively, since the session and matrix are saved together, the session can be recalled from the matrix.

## Case Listing Data Tab

On the Case Listing Data Tab you will select the database containing the records you wish to view. You may choose an Incidence, Population, Standard Million, or Expected Rates database.

## Database for a Case Listing Session

From the Case Listing Data Tab, click on the incidence, population, or expected rates database containing the records you wish to view. It is important to choose the database first in a Case Listing session. Each database has a unique data dictionary which lists the available variables and their values. These variables are used throughout the session.

### Using Your Own Data

If you would like to use your own data with SEER\*Stat, it must be processed using SEER\*Prep. SEER\*Prep is available on the SEER Program's web site under the topic of Scientific Systems:

[www.seer.cancer.gov](http://www.seer.cancer.gov)

## Case Listing Selection Tab

When creating a Case Listing Session, the search statements on the Case Listing Selection Tab define for SEER\*Stat the subset of cases to be viewed. The variables available are dependent on the database you selected on the Data tab.

Search Statements are created on the Case Selection Window. To access the Selection Window click the Edit button next to the text box. If there are no search statements on the Selection window, then all cases in the database will be selected. Therefore, you should only make search statements for variables used to exclude cases.

To clear all search statements from the Selection box, click the Clear button.

For further instructions on making selections see Working with Search Statements and Tips for Accurate Search Statements.

## Case Listing Table Tab

On the Case Listing Table Tab you may select the column or sort variables for your output matrix. It is not possible to execute a Case Listing Session without choosing a least one column variable.

### Column and Sort Lists

These are lists of variables for the column or sort order of your output matrix.

### Remove Buttons

Click  to remove a table variable.

### Up and Dn Buttons

Click  or  to change the order of table variables.

### Variables

The variables in the database are listed in categories. Each category is named for the variables found in it. To expand a category, double click the category label or click the plus sign (+) next to it. This will reveal the variables contained in the category. Single click a variable name to select it. To view the definition of a variable in the dictionary, double click the variable name and click the Create or Modify buttons on the Dictionary window.

### Column, Sort, or Both Buttons

Click these buttons to add the selected variable to the column, sort, or both lists. You may add all the variables from a category at once by selecting the category and clicking the Column, Sort, or Both buttons.

### Find Button

If you are looking for a particular variable and cannot find it within the variable folders, click the Find button to bring up the Find Window.

## Add a Table Variable to a Case Listing Session

- 1 Click on a variable in the Variable section of the Case Listing Table Tab or click the Find button to select the variable you want. You can add all the variables in a category by selecting the category name.

- 2 Click the Column, Sort, or Both buttons. The name of the variable will appear in the appropriate lists. There is a limit of 50 sort variables you may add to the list.

## Order the Table Variables in a Case Listing Session

- 1 From the Case Listing Table Tab click the variable you want to move within the list (Column or Sort).
- 2 Use ↑ or ↓, the Move Up and Move Down commands from the Edit menu, or Ctrl + Up Arrow and Ctrl + Down Arrow on the keyboard to move the variable up or down on the list.

Variables may be moved only within a single list. To move a variable from one list to another you must remove the variable from one list and add it to the other.

## Remove a Table Variable from a Case Listing Session


- 1 From the Case Listing Table Tab, click the variable you want to remove from the Column or Sort list.
- 2 Click the × button, select Delete/Remove from the Edit menu, or press the Delete key on your keyboard.

## Case Listing Output Tab

On the Case Listing Output Tab you can add a title to your output matrix. Simply type your title in the text box provided. When the matrix is printed, the text will be centered on the page, but will not wrap to a second line without carriage returns. To get a multi-lined title, press Enter at the end of each line of text.

## Title for a Case Listing Matrix

If you wish to add a title to your output matrix, simply type the title in the text box provided on the Case Listing Output Tab. When the matrix is printed, the text will be centered on the page, but will not wrap to a second line without carriage returns. To get a multi-lined title, press Enter at the end of each line of text.

Once the matrix has been made you can see the entire title by clicking on  or choosing Options from the Matrix menu. The title may be added or edited on the Case Listing Matrix Options Window. Only the first line of the title will be displayed on the SEER\*Stat matrix window. To see how the Matrix title will print, select Print Preview from the File menu.

# Dictionary


The dictionary contains the definitions of all available variables in the selected database. Initially, the dictionary contains only the definitions of the standard database variables. As you work with the database, you will have the opportunity to create user-defined and merged variables which will also appear in the dictionary.

When you open the dictionary it is actually the dictionary for the currently selected database on the Data Tab. While a session is open, the dictionary of the currently selected database will not change based on actions taken in other open sessions. However, if you select a different database on the Data Tab and then return to the originally selected database, the dictionary will be updated.

## Note

It is extremely important to select the database you want from the Data Tab of your session before accessing the Dictionary. Otherwise, you might put a lot of work into creating a user-defined or merged variable for your session, only to discover that it is based on variable(s) not found in the database you actually want to use. Your new variable will not be available unless you export it from the database in which you created it and import into the database in which you want to use it.

## Open the Dictionary

- 1 Open a Frequency/Rate Session, Survival Session, or a Case Listing Session.
- 2 Select a database from the Data Tab.
- 3 Click  or select Dictionary from the File menu to access the dictionary.

From the dictionary, you may view variables in the dictionary and create user-defined and merged variables. Merged variables may only be created in Frequency/Rate and Survival Sessions.

## Dictionary Window

The dictionary contains the definitions of all available variables associated with the currently selected database.

### Categories and Variables

All variables are arranged into categories. A category is named for the variables found in it. To expand a category, double click the category label or click the plus sign (+) next to it. This will reveal the variables

contained in the category. The following icons are used to designate the type of variable in the database.



represents a standard variable



represents a user-defined or merged variable created for the session, but not saved in the dictionary



represents a user-defined or merged variable saved in the database dictionary

The definition of each variable is held in the dictionary. To see the definition of a particular variable, click the variable name and click the Create button.

### **Create Button**

Allows you to create a new user-defined variable or merged variable based on the currently selected variable. If the currently selected variable is a standard variable or user-defined variable, clicking the Create button will open the Edit Variable Window for a new user-defined variable. If the currently selected variable is a merged variable, clicking the Create button will open the Merged Variable Window for a new merged variable.

### **Modify Button**

Allows you to modify the currently selected user-defined variable by sending you to the Edit Variable Window or merged variable by sending you to the Merged Variable Window.

### **Delete Button**

Removes the currently selected user-defined or merged variable from the dictionary.

### **Merge Button**

Allows you to create a new merged variable by sending you to the Merged Variable Window. There you can create a variable which is based on more than one standard or user-defined variable. This button is disabled in Case Listing Sessions.

### **Import Button**

Allows you to add user-defined and merged variables which have been exported from a different database to the current database dictionary as either temporary or permanent variable definitions. Select the file you wish to import on the Import User-Defined Variables Window and press OK. The Import User Variable Definitions Window will appear. Here you may select the variables which you want to import.

### **Export Button**

Allows to you save a file containing user-defined and merged variables definitions which may be applied to another database or shared with




another SEER\*Stat user. When the Export User Variable Definitions Window appears, enter the directory path and name for your new export file. If you do not know the entire path name, click the Browse button to select the path. Select the user-defined and merged variables you wish to be available to other database dictionaries and click OK.

### **Find Button**

If you are looking for a particular variable, but cannot find it within the variable folders, click here to bring up the Find Window.

## **Import User Variables**

To add user-defined or merged variables which were prepared for a different database or by another user into your current database dictionary, follow these steps.

- 1 Click  or select Dictionary from the File menu to access the currently selected database dictionary.
- 2 Click Import. The Import User Variable Window will appear.
- 3 Select the file which contains the variables you will be importing and click OK. The Import User Variable Definitions Window will appear.
- 4 Select the variables you wish to import.
- 5 Confirm that the Save to Dictionary check box is marked as you wish.
- 6 Click OK. The newly imported variables will appear in the User-Defined or Merged variable categories.

### **Note**

Merged variables may only be imported for use in Frequency/Rate and Survival Sessions.

## **Import User Variable Definitions Window**

Access the Import User Variable Definitions Window by clicking the Import button on the Dictionary Window. Initially, the Import User Variable Window will appear, which is where you will select the directory path and filename of the file to be imported into the current database dictionary. Once selected, click Open and the Import User Variable Definitions Window will appear.

### **Filename**

This is the directory path and filename selected on the Import User Variable Window. This field is only for display. If you need to change the path and filename, click Cancel and begin again.

### **Variables**

Select the variables you wish to import to the current database dictionary. Variable definitions which depend on standard variables that are either different or not available in the current database will not be imported. It is possible that the name of a variable being imported and the name of an existing variable in the current database dictionary could be the same, but have different definitions. Should this happen, you will be given the option to continue to import the new variable or to keep the old as it is.

### **Select All Button**

Allows you to select all variables at once.

### **Select None Button**


Allows you to deselect all variables at once.

### **Save to Dictionary**

When checked, all imported variables will be saved to the current database dictionary for use in this and future sessions. Once a user-defined variable is saved to the database dictionary, all updates and modifications will be saved to the database as well.

## **Export User Variables**

To create an export file so that you may share user-defined or merged variables across databases or with another user, follow these steps.

- 1 Click  or select Dictionary from the File menu to access the currently selected database dictionary.
- 2 Click Export. The Export User Variable Definitions Window will appear.
- 3 Give the directory path and filename for the export file. The file extension .fmx must be used for all export variable files.
- 4 Select the variables you wish to export from the list of all user-defined and merged variables from the current database dictionary.
- 5 Click OK.

### **Note**

Merged variables may only be exported for use in Frequency/Rate and Survival Sessions.

## **Export User Variable Definitions Window**

Access the Export User Variable Definitions Window by clicking the Export button on the Dictionary Window.

**Filename**

Type the directory path and filename for your new export variable file. The file extension .fmx is required for all export variable files.

**Variables**

Select the variables which you wish to export from the list of user-defined and merged variables found in the current database dictionary.

**Select All Button**

Allows you to select all variables displayed at once.

**Select None Button**

Allows you to deselect all variables displayed at once.

## Database Variables

All variables are arranged into categories. Most categories are named for the variable types found in it. For example, if you are using the SEER Public-Use data, all of the Extent of Disease variables are in their own category.

For incidence and population databases, the first two categories in the list were created specifically for Frequency/Rate statistics, but for consistency have been made standard throughout SEER\*Stat. These categories contain the variables which are found in the Standard Million and Population files. It is a good idea to check these two categories first before going lower in the list when you are searching for a particular variable.

The User-Defined and Merged variables categories are listed last, when available. You can edit any standard or user-defined variable and preserve the modification as a user-defined variable. You may also create merged variables by defining groupings which are based on the values of one or more variables. When you create a user-defined or merged variable, it is important to give the variable a meaningful name since they will be used as row, column, and page identifiers in the output matrix. The following icons are used to designate the type of variable in the database.



represents a standard variable



represents a user-defined or merged variable created for the session, but not saved in the dictionary




represents a user-defined or merged variable saved in the database dictionary

**Note**

Merged variables may only be created in Frequency/Rate or Survival sessions.

## Standard Variables

Each SEER\*Stat database has its own list of standard variables. To view the definition of a standard variable, follow these steps:


- 1 Open the dictionary by clicking  or selecting Dictionary from the File menu.
- 2 Select the variable you want to view. You may use the Find Button, if desired.
- 3 Click the Create button. This will open the Edit Variable Window containing the variable you selected.
- 4 To exit the Edit Variable Window without creating a user-defined variable, click the Cancel button.


## User-Defined Variables

A user-defined variable is an edited version of one standard variable from the current database. You should create a user-defined variable when you want to change the organization of one variable's values for display in an output matrix. Create them in order to:

- Combine values into new groupings.
- Change the values associated with the groupings.
- Rename existing groupings in the variable.
- Reorder existing groupings.
- Remove existing groupings from the variable.

User-defined variables are always saved with the session in which they were created and used, but can also be saved in the database dictionary for use in later sessions. Saved user-defined variables are written to disk with file extension '.fmt'. Do not delete or edit the '.fmt' file. On the Preferences Window, you may set the default to check or uncheck the Save to Dictionary option on the Edit Variable Window. The following icons are used to designate the type of variable in the database.

 represents a user-defined or merged variable created for the session, but not saved in the dictionary


 represents a user-defined or merged variable saved in the database dictionary


When you open the dictionary it is actually the dictionary for the currently selected database on the Data Tab. When opening a previously saved session that contains a user-defined variable not saved to the dictionary, the session-specific user-defined variable will be replaced by a database-wide variable if they possess the same variable name.

## Merged Variables

A merged variable is a user created variable with groupings defined from specified values from one or more variables. You should create a merged variable when you want to use the values of two or more standard variables to define the groupings. Merged variables cannot be created for use in a Case Listing Session.

Merged variables are always saved with the Frequency/Rate or Survival session in which they were created and used, but can also be saved in the dictionary for use in later sessions. Saved merged variables are written to disk with file extension '.fmt'. Do not delete or edit the '.fmt' files. On the Preferences Window, you may set the default to check or uncheck the Save to Dictionary option on the Merged Variable Window. The following icons are used to designate the type of variable in the database.

 represents a user-defined or merged variable created for the session, but not saved in the dictionary


 represents a user-defined or merged variable saved in the database dictionary

When you open the dictionary it is actually the dictionary for the currently selected database on the Data Tab. When opening a previously saved session that contains a merged variable not saved to the dictionary, the session-specific merged variable will be replaced by a database-wide variable if they possess the same variable name.

## Working With Variables


To view the definition of any database variable, to create, modify, or delete a user-defined or merged variable, or to import and export user-defined and merged variables, you must access the Dictionary Window. Please see Opening the Dictionary for instructions.

### Create a User-Defined Variable


- 1 Open the Dictionary by clicking .
- 2 Expand the categories until you find the variable on which you will base your user-defined variable or click the Find button to use the Find Window. You may create a user-defined variable from either a standard variable or a user-defined variable.
- 3 Click on this variable.
- 4 Click Create. The Edit Variable Window will appear.
- 5 Edit the variable.

- 6 Check Save to Dictionary if you want the variable available for database-wide use. On the Preferences Window, you may set the default to check or uncheck the Save to Dictionary option on the Edit Variable Window.
- 7 Click OK.


## Create a Merged Variable

- 1 Open the Dictionary by clicking  .
- 2 You may either create a merged variable from scratch or create it from the definition of an existing merged variable in a Frequency/Rate or Survival session.
  - A. To create a new merged variable from scratch, click the Merge button. A blank Merged Variable Window will appear.
  - B. To create a new merged variable based on an existing merged variable, select the existing merged variable from the Merged category on the Dictionary Window and click the Create button. The groupings defined for the selected merged variable will be shown on the Merged Variable Window that appears.
- 3 Edit the variable.
- 4 Check Save to Dictionary if you want the variable available for database-wide use. On the Preferences Window, you may set the default to check or uncheck the Save to Dictionary option on the Merged Variable Window.
- 5 Click OK.


## Modify a User-Defined Variable

- 1 Open the Dictionary by clicking  .
- 2 Find the user-defined variable you wish to modify.
- 3 Click on this variable.
- 4 Click the Modify button. The Edit Variable Window will appear.
- 5 Edit the variable.
- 6 Check Save to Dictionary if you want the variable available for database-wide use. If the variable has already been saved for database-wide use, then the Save to Dictionary checkbox is permanently checked.
- 7 Click OK.

## Modify a Merged Variable

- 1 Open the Dictionary by clicking .
- 2 Find the merged variable you wish to modify for a Frequency/Rate or Survival session.
- 3 Click on this variable.
- 4 Click the Modify button. The Merged Variable Window will appear.
- 5 Edit the variable.
- 6 Check Save to Dictionary if you want the variable available for database-wide use. If the variable has already been saved for database-wide use, then the Save to Dictionary checkbox is permanently checked.
- 7 Click OK.

## Delete User-Defined or Merged Variables

- 1 Open the Dictionary by clicking .
- 2 Find the user-defined or merged variable you wish to delete. You may delete only user-defined and merged variables.
- 3 Select the variable.
- 4 Click the Delete button. The variable will be deleted.

## Save a User-Defined Variable to the Database Dictionary

Sometimes it is beneficial to save a commonly needed user-defined variable to the database dictionary for use in other sessions. The decision to use a user-defined variable in the current session only or to save it to the database dictionary is made on the Edit Variable Window. Once a user-defined variable is saved to the dictionary, all updates and modifications will be saved to the dictionary as well. On the Preferences Window, you may set the default to check or uncheck the Save to Dictionary option on the Edit Variable Window.

- 1 Check Save to Dictionary only if you want to save the user-defined variable for database-wide use.
- 2 Click OK.

## Save a Merged Variable to the Database Dictionary


Sometimes it is beneficial to save a commonly needed merged variable to the database dictionary for use in other Frequency/Rate or Survival sessions. The decision to use a merged variable in the current session only or to save it to the database dictionary is made on the Merged Variable Window. Once a merged variable is saved to the database dictionary, all updates and modifications will be saved to the dictionary as well. On the Preferences Window, you may set the default to check or uncheck the Save to Dictionary option on the Merged Variable Window.

- 1 Check Save to Dictionary only if you want to save the user-defined variable for database-wide use.
- 2 Click OK.

## Edit Variable Window

### Open the Edit Variable Window

Access the Edit Variable Window when you want to create or modify a user-defined variable. It can also be used to view the definition of a standard variable or an existing user-defined variable. To access the Edit Variable Window:

- 1 Open the dictionary by clicking .
- 2 Select a standard or user-defined variable.
- 3 Click the Create or Modify buttons. The Edit Variable Window for the variable you selected will appear. You cannot modify a standard variable.

### Edit Variable Window

On the Edit Variable Window you can create or modify a user-defined variable by changing the groupings of an existing variable.

#### Name

Every variable in the dictionary must have a unique name. It is a good idea to give meaningful names to new user-defined variables so that they may be easily identified.

#### Description

A short description about the standard variable's values may appear in this field. In some cases, hyperlinks to documentation are also included. If you have an internet browser installed and are connected to the Internet, you may click the hyperlink for immediate access to this documentation.



**Grouping**

Each variable is defined by groupings of its values. The groupings need not cover all possible values and may overlap when appropriate.

**Values**

Every variable has a list of all possible entries, called the values. Most values have descriptive labels for discrete groupings. The values available in a variable cannot be changed. Variables are defined and modified through the groupings of the values.

**Unlabeled Values**

There are a few variables with unlabeled values. These values are represented by a range of numbers. Hyphens ( - ) and commas ( , ) may be used to define/specify a range, for example: 2,5-7,9. There is one exception to the rule that all ranges are numeric in the SEER Public-Use data. The Recode ICD-O-2 to 9 variable also includes an underscore in the right-most position of an unlabeled value which represents a blank. For example, 140\_-1409 is the range of all values starting with 140 in this particular variable.

**Add Button**

Used to create a new grouping of values.

**Update Button**

Used to change the selection of values assigned to an existing grouping.

**Add All Button**

Used to add all values to the variable definition either individually as separate groupings or as a single grouping. The individual grouping labels are the same as the value names. The combined values grouping is labeled "All Values" by default.

**Add Rest Button**

Used to add all values, which did not fall into one of the groupings already specified, to the variable definition either individually as separate groupings or as a single grouping. The combined values grouping is labeled "Remaining Values" by default.

**Up/Dn Buttons**

Changes the order of existing groupings without affecting the values assigned to them. The order of groupings in the output matrix is determined by the order of groupings in this window.

**Rename Button**

Allows you to rename an existing grouping without affecting the values assigned to the grouping. It is important to give meaningful names to the

groupings since they will be used as the row, column, and page identifiers in the output matrix.

### **Delete Button**

Deletes the selected grouping(s).

### **Save to Dictionary Check Box**

When checked, the user-defined variable is saved to the dictionary for use in this and future sessions. Once a user-defined variable is saved to the dictionary, all updates and modifications will be saved to the database as well. On the Preferences Window, you may set the default to check or uncheck the Save to Dictionary option.

### **Add a New Grouping to a User-Defined Variable**

The Add button on the Edit Variable Window is used to create a new grouping for a user-defined variable.

- 1 Highlight the values you wish to include in the grouping.
- 2 Click the Add button.
- 3 Type a suitable label for the new grouping, since this is the label that will appear in your output matrix.
- 4 Check the values assigned to the grouping for accuracy.

You can select more than one value using any of these methods.

- Hold down the Control key while you click on values you want to select or deselect. Every value you select will be highlighted.
- Hold down the Shift key while you click the first and last values of a list you want to select. The two values you select and all values in between will be highlighted.
- Click and drag over the values you want to select.

### **Add All Variable Values to a User-Defined Variable**

The Add All button on the Edit Variable Window is used to add all values to the variable definition either individually as separate groupings or as a single grouping. Beware that using this option may create over-lapping groupings in your variable.

- 1 When you are ready to add all values, click the Add All button. A dialog box will appear with three choices: Combined, Individual, and Cancel.
  - The Combined option will place all values into one grouping with the default label "All Values."

- The Individual option will add each of the values as separate groupings where the grouping labels are the same as the value names.
- 2 Choose the option you want.
  - 3 Rename any groupings as needed.
  - 4 Check the values assigned to the groupings for accuracy.

### **Add Rest of the Variable Values to a User-Defined Variable**

The Add Rest button on the Edit Variable Window is used to add all values that do not fall into an existing grouping to the variable definition, either individually or as a single grouping.

- 1 When you are ready to add the rest of the values, click the Add Rest button. A dialog box will appear with three choices: Combined, Individual, and Cancel.
  - The Combined option will place any unused values into one grouping with the default label “Remaining Values.”
  - The Individual option will add the rest of the values as discrete groupings where the grouping labels are the same as the value names.
- 2 Choose the option you want.
- 3 Rename any groupings as appropriate.
- 4 Check the values assigned to the groupings for accuracy.

### **Update the Definition of an Existing Grouping in a User-Defined Variable**

The Update button on the Edit Variable Window is used to change the values assigned to an existing grouping.

- 1 Highlight the grouping you want to update.
- 2 Highlight the values you wish to include in the grouping.

You can select more than one value using any of these methods.

  - Hold down the Control key while you click on values you want to select or deselect. Every value you select will be highlighted.
  - Hold down the Shift key while you click the first and last values of a list you want to select. The two values you select and all values in between will be highlighted.
  - Click and drag over the values you want to select.
- 3 Click the Update button.
- 4 Check the values assigned to the grouping for accuracy.

## Rename an Existing Grouping in a User-Defined Variable

You may rename an existing grouping on the Edit Variable Window without affecting the values assigned to the grouping. It is important to give meaningful names to the groupings since they will be used as the row, column, and page identifiers in the output matrix.

- 1 Click the grouping label you want to rename to highlight it.
- 2 Click the Rename button or click the grouping label again.
- 3 Type the new name for the grouping and press Enter.

## Order the Groupings of a User-Defined Variable

You may change the order of existing groupings on the Edit Variable Window without affecting the values assigned to them. The order of the groupings in the output matrix is determined by the order of the groupings in this window.

- 1 Click the grouping(s) you want to move.
- 2 Click the Up or Dn buttons.

If non-consecutive groupings are selected, they will move together with the most outer grouping of those selected determining what up and down mean. For example, suppose you wanted to change the order of SEER variable Race Recode A.

All Races  
White  
Black  
Other  
Unknown

If you select the groupings *All Races* and *Black*, then only the Dn button will be enabled, because *All Races* cannot go any higher (Up) in the list of groupings. When the Dn button is clicked, the order will change so that *Black* will drop beneath *Other* and *All Races* will join *Black* there.

White  
Other  
All Races  
Black  
Unknown

## Delete an Existing Grouping from a User-Defined Variable

You may delete existing groupings on the Edit Variable Window.

- 1 Click the groupings you want to remove.

You may select more than one grouping using any of these methods.


- Hold down the Control key while you click on groupings you want to select or deselect. Every grouping you selected will be highlighted.
- Hold down the Shift key while you click the first and last labels of a range of groupings you want to select. The two groupings you selected and all groupings in between will be highlighted.
- Click and hold to the right of a grouping you want to select. Drag the Lasso to the left and up or down to highlight the groupings.

- 2 Click the Delete button or press the Delete key on your keyboard.

## Merged Variable Window

### Open the Merged Variable Window

Access the Merged Variable Window when you want to view the definition of an existing merged variable, create a new merged variable, or modify an existing merged variable.

- 1 Open the dictionary by clicking .
- 2 To create a new merged variable, select a standard or user-defined variable and click the Merge button.
- 3 To edit an existing merged variable click the Create, or Modify buttons.

The Merged Variable Window for the variable you selected will appear. If the Create or Modify buttons were clicked, the Merged Variable Window will contain the definition of the selected merged variable. If the Merge button was clicked, the Merged Variable Window will be blank. The Merge button is disabled in a Case Listing session.

### Note

You may only create merged variables for use in Frequency/Rate or Survival sessions.

### Merged Variable Window

On the Merged Variable Window, you have the opportunity to create a variable with groupings based on one or more variables. Only standard or user-defined variables can be used to create merged variables, a merged variable cannot be based on other merged variables.

### Name

Every variable in the Dictionary must have a unique name. It is a good idea to give a meaningful name to a merged variable so that it may be easily recognized in the list of database variables.

## **Groupings**

The names of defined groupings are listed here. It is important to give meaningful names to the groupings since they will be used as the row, column, or page identifiers in the output matrix.

### **Add Button**

Used to define a new grouping. When you click the Add button, the New Merged Grouping Window (Selection Window) appears.

### **Add All Button**

Used to define all the possible grouping combinations from one or more variables. When you click the Add All button, the Add All Merged Grouping Window appears.

### **Edit Button**

Used to check or change the definitions of an existing grouping. Select the grouping you wish to view or edit and click the Edit button. The existing definition of the grouping will appear in the text box at the top of the Edit Merged Grouping Window (Selection Window). Notice that the grouping name appears in the title bar of the Edit Merged Grouping Window.

### **Delete Button**

Deletes the selected grouping(s).

### **Rename Button**

Allows you to rename an existing grouping without affecting the definition of the grouping.

### **Move Up/Move Dn Buttons**

Changes the order of existing groupings without affecting the definitions of the groupings. The order of the groupings in the output matrix is determined by the order of the groupings in this window

### **Grouping Definition**

Displays the search statement that defines the selected grouping. Text may be copied from this box, but to make changes to a grouping definition, you must highlight the grouping and click the Edit button to open the Edit Merged Grouping Window (Selection Window).

### **Include “Other” Grouping Check Box**

Check this box to ensure that no cases are excluded in the output matrix. When checked, cases that do not fall into a defined grouping will automatically be added to a grouping labeled “Other” in the output matrix. It is not possible to change the label of the “Other” grouping.

### **Save to Dictionary Check Box**

When checked, the merged variable is saved to the dictionary for use in this and future sessions. Once a merged variable is saved to the dictionary, all updates and modifications will be saved to the dictionary as well. On the Preferences Window, you may set the default to check or uncheck the Save to Dictionary option.

### **Add a New Grouping to a Merged Variable**

- 1 From the Merged Variable Window, click the Add button. The New Merged Grouping Window (Selection Window) will appear.
- 2 On the New Merged Variable Window (Selection Window), create a search statement for the subset of cases the new grouping will contain. Notice that the Not button is disabled and the only operator available is the “is = to.”
- 3 Once the search statement is complete, click OK to return to the Merged Variable Window.
- 4 Type in a meaningful name for the grouping since this will be the label that will appear in your output matrix.

### **Note**

Be careful to consider adding the sex variable to any definition which contains a sex specific cancer site. For example, suppose you wanted a grouping which includes both female breast cancer cases and cervical cancer cases. You might write the search statement as follows:

(sex = female and site = breast) or site = cervix

However, when applying a merged variable grouping definition to the population file for the creation of rates or trends, any variable in the definition which is not found in the population file is evaluated as true. In this example, our statement would be evaluates like this:

(sex = female and site = breast) or site = cervix

(sex = female and TRUE) or TRUE

sex = female or TRUE

TRUE

Notice that the population for the denominator of this rate will contain both males and females. To correct this problem, you might change the search statement as follows:

sex = female and (site = breast or site = cervix)

This new search statement will be evaluated for the population file subset in this way:

sex = female and (site = breast or site = cervix)

sex = female and (TRUE or TRUE)



sex = female and TRUE

sex = female

This is what should be the denominator for a rate of two female sites.

## **Add All Groupings to a Merged Variable**

This is used to define all the possible grouping combinations from one or more variables (only standard or user-defined variables can be used to create merged variables, a merged variable cannot be based on other merged variables).

- 1 From the Merged Variable Window, click the Add All button. The Add All Merged Groupings Window will appear.
- 2 On the Add All Merged Groupings Window, select one or more variables from the Available Variables box and use the arrow buttons,  or , to move them into the Selected Variables box.
- 3 Click OK to return to the Merged Variable Window.
- 4 All the possible groupings that can be created by combining the chosen variables will be displayed. You may use the Rename button to assign your own labels for the groupings.

Suppose you wanted a merged variable based on all the combinations of the Sex and the Race Recode A variables. If you follow the steps above, when you click the OK button on the Add All Merged Groupings Window, the Merged Variable Window would contain the following groupings:

Male and female/All races

Male and female/White

Male and female/Black

Male and female/Other

Male and female/Unknown

Male/All races

Male/White

Male/Black

Male/Other

Male/Unknown

Female/All races

Female/White

Female/Black



Female/Other

Female/Unknown

## **Add All Merged Groupings Window**

On the Add All Merged Groupings Window, you can define a merged variable based on all possible groupings that can be created by combining the groupings from one or more variables.

### **Available Variables**

This box contains a list of all the standard and user-defined variables in the database arranged into categories. To expand a category and view the variables, double click the category label or single click the plus sign (+) next to it.

### **Selected Variables**

This box contains the database variables you choose to define the groupings for the merged variables. To choose a variable, select a variable from the Available Variables box and use the arrows between the two boxes to move it to the Selected Variables box.

### **Move Up/Move Dn Buttons**

These buttons allow you to change the order of the selected variables, which will change the order of the grouping labels in the variable without changing the grouping definitions.

## **Edit a Grouping Definition in a Merged Variable**

- 1 From the Merged Variable Window, select the grouping you wish to edit.
- 2 Click the Edit button. The Edit Merged Grouping Window (Selection Window) will appear with the grouping name in the Window title bar.
- 3 The easiest way to edit a search statement is to select (or highlight) the entire search statement in the text box at the top of the window and hit the Delete key on your keyboard. This will remove the entire search statement. However, if you only need to make a small change to a complicated search statement, consult the following topics for specific instructions.

Edit a Search Line or Changing the Conjunction

Edit Parentheses in a Search Statement

Delete a Search Line from a Search Statement

Order the Lines in a Search Statement

- 4 When you are finished editing the grouping, click OK to return to the Merged Variable Window.

## Delete a Grouping from a Merged Variable

- 1 From the Merged Variable Window, select the grouping(s) you want to delete.
- 2 Click the Delete button.

## Rename a Grouping in a Merged Variable

It is important to enter meaningful names for your groupings since they will be used as row, column, and page identifiers in the output matrix.

- 1 From the Merged Variable Window, select the grouping you want to rename
- 2 Click the Rename button or click again on the grouping name.
- 3 Type the new name.
- 4 Hit Enter on your keyboard or click away from the grouping's name.

## Order the Groupings of a Merged Variable

You may change the order of existing groupings on the Merged Variable Window without affect the values assigned to them. The order of the groupings in the output matrix is determined by the order of groupings on this window.

- 1 From the Merged Variable Window, select the grouping(s) you wish to move up or move down in the list.
- 2 Click the Move Up or Move Dn buttons as desired.

If non-consecutive groupings are selected, they will move together with the most outer grouping of those selected determining what up and down mean. For example, suppose you wanted to change the order of a merged variable with the following groupings.

White Male  
White Female  
Black Male  
Black Female  
Other race Male  
Other race Female

If you select the groupings *White Male*, *Black Male*, and *Other race Male*, then only the Dn button will be enabled, because the *White Male* grouping cannot go any higher (Up) in the list of groupings. When the Dn button is clicked, the order will change so that *Other race Male* will drop beneath *Other race Female* and the other two selected groupings will join *Other race Male*.

White Female  
Black Female  
Other race Female  
White Male  
Black Male  
Other race Male

### **Include “Other” Grouping in a Merged Variable**

When this box is checked on the Merged Variable Window, any cases that do not fall into a defined grouping are automatically added to a grouping labeled “Other” in the output matrix. When creating complex merged variables, it can be difficult to tell whether all cases will be captured in the defined groupings. It is not possible to change the label of the “Other” grouping because the excluded cases are not determined until the job has finished executing.



# Working with Search Statements

## Search Statements

Search statements are used to define a subset of records in the database to select cases for a session or for a grouping in a merged variable. When creating a session, a search statement on the Selection Tab tells SEER\*Stat which subset of cases you wish to use in your analysis. Click the Edit button on the Selection Tab to access the Selection Window where search statements are created. When creating a merged variable, a search statement serves as the definition for each grouping in the variable. Click the Add or Edit buttons on the Merged Variable Window to access the Merged Grouping Window (Selection Window).

## Selection Window

On the Selection Window (Merged Grouping Window) you will build the search statement which defines the subset of cases for your session or a grouping for your merged variable.

### Search Statements

The basic building block of a search statement is the search line. A search line must consist of the variable name, an operator, and the variable groupings to be selected or excluded. To define additional search lines, click the And or Or buttons before selecting another variable. You may also add parentheses ( .. ) and Not to the search statement for additional meaning.

### Variable

The Variable box contains the names of all the standard variables in the Dictionary. When a variable name is selected, it will appear in the Search Statements box at the top of the Selection Window.

### Operator

There are two equality operators: Equal Sign or Not Equal Sign. Choose “is = to” to select and “is not = to” to exclude the variable groupings from the analysis. When creating a merged variable grouping, only the “is = to” operator is available.

### Selection

The Selection(s) box contains a list of all groupings and the range of continuous values defined for the selected variable. Only the groupings defined for the variable will be listed. Identify the groupings you wish to select by clicking on them, or for continuous values entering the range in the text box using hyphens ( - ) and commas ( , ). To select more than one grouping at a time use one of these methods.

- Hold down the Control key while you click on groupings you want to select or deselect.
- Hold down the Shift key while you click the first and last groupings of a list you want to select. The two groupings you selected and all groupings in between will be highlighted.
- Click and drag over the groupings you want to select.

### **( .. ) Button**

Parentheses around a group of search lines tells the computer to implement those search lines first when evaluating the search statement. Therefore it is a good idea to add parentheses when more than two search lines are needed to clarify the search statement.

### **Not Button**

Not Tells the computer to select the subset which is the opposite of the group described after the Not. When creating a merged variable grouping, this button is disabled.

### **And Button**

When adding a new search line to the statement, a conjunction is required. Only one may be selected at a time. And tells SEER\*Stat to select the subset of cases where both search lines on either side are true.

### **Or Button**

When adding a new search line to the statement, a conjunction is required. Only one may be selected at a time. Or tells SEER\*Stat to select the subset of cases where one of the search lines on either side of the Or is true.

### **Precedence in a Search Statement**

The precedence of the Boolean operators is indicated by the order they are listed below. That is, parentheses have the highest precedence and Or has the lowest precedence. Precedence for multiple instances of the same operator is determined by the order it appears in the statement (first to last).

(. . )

Not

And

Or

### **Find Button**

If you are looking for a particular variable, but cannot find it within the variable folders, click here to bring up the Find Window.

## Accessing the Selection Window

The Selection Window (Merged Grouping Window) can be accessed while setting up a Frequency/Rate Session, Survival Session, Case Listing Session, or when creating a merged variable in the Dictionary.

- To access the Selection Window from a session, where you may build your search statement, click on an Edit button found on the session Selection Tab.
- To access the Merged Grouping Window while creating a merged variable, click the Add or Edit buttons on the Merged Variable Window.

## Building Search Statements with Search Lines

The basic building blocks of a search statement are search lines. A search line must consist of the following items:

- 1 Variable Name -- chosen from the Variable box
- 2 “is = to” or “is not = to”—chosen from the Operator box
- 3 Variable groupings—chosen from the Selection(s) box

Note that when you are working with a merged variable, the “is not = to” operator is disabled on the Merged Grouping Window (Selection Window).

## Create a Search Line

- 1 Find and click the variable name in the Variable box which contains all the standard and user-defined variable names in the Dictionary. When the variable name is selected, it will appear in the Search Statements box at the top of the Selection Window. If you have difficulty locating the variable name, click the Find button to bring up the Find Window.
- 2 Click on either “is = to” or “is not = to” in the Operator box; either an “=” or “!=” will appear. “!=” means “not equal.”
- 3 Select the variable groupings from the Selection(s) box you wish to include (or exclude) from your case subset or grouping definition. You will either be adding groupings, continuous values, or both to the search line.
- 4 To define additional search lines, click the And or Or buttons and repeat steps 1 through 3. You may also add parentheses ( .. ) and Not to the search statement for additional meaning.
- 5 When all of your search lines have been added to the search statement, click OK.

### Note

When defining a grouping for a merged variable, the “is not = to” operator and the Not button are unavailable on the Merged Grouping Window.

## Edit a Search Line

If the wrong variable name, operator, or selections were chosen in your search statement, follow these steps.

- 1 Place your cursor anywhere in the search line you wish to edit. The current variable name, operator, and grouping selections for the line will be displayed in the Variable, Operator, and Selection(s) boxes.
- 2 Click the correct variable name, operator, and make the appropriate selections.

## Find Window

The Find Window allows you to find variable names easily from the Selection Window, Table Tab, or Dictionary.

- 1 Type the first three letters of the variable name.  
SEER\*Stat will search for that combination of letters anywhere in the names of the database variables. For example, if you are using the SEER Public-Use database, you might type “age” to find age variables. Not only will SEER\*Stat display the age-related variables, but the “stage” variables will also be listed.
- 2 Select the variable name you want and click OK.  
The variable you selected will be highlighted for you.

## Selection(s) Box on the Selection Window

The Selection(s) box contains a list of all variable groupings and the range of continuous values defined for the variable you chose for your search statement. Variables may have groupings only, continuous values only, or both groupings and continuous values.

- The groupings are listed in the scrolling list box. Only the groupings defined for the variable will be listed.
- The range of continuous values is listed to the right of the Selection(s) title. This range is for your information and cannot be selected. The single line text box under the title is where you can select the continuous values by typing the range of values you want using hyphens ( - ) and commas ( , ). For example “2,5-7,9” selects the values 2, 5, 6, 7, and 9.

## Select Groupings for a Search Line

Any number of groupings may be selected for your search statement. Once a grouping is selected, it is displayed in the search statement at the top of the Selection Window and highlighted in the Selection(s) list box.

- To select a single grouping, click the grouping label.



- If the wrong grouping is chosen, click the correct grouping.

You can select more than one grouping using any of these methods.

- Hold down the Control key while you click on groupings you want to select or deselect.
- Hold down the Shift key while you click the first and last groupings of a list you want to select. The two groupings you selected and all groupings in between will be highlighted.
- Click and drag over the groupings you want to select.

## **Select Continuous Values for a Search Line**

There are a few variables with unlabeled values. These values are represented by a range of numbers. To select continuous values of a variable for your search statement, they must be entered in the Selection(s) text box. Once the values are entered, press the Enter key to add the values to the search line. Enter the values in the text box using the following rules.

- Use a hyphen ( - ) to specify a range of values: 10-25
- Use a comma ( , ) to separate ranges or single values: 44-56,78,80-99

There is one exception to the rule that all ranges are numeric in the SEER Public-Use data. The Recode ICD-O-2 to 9 variable also includes an underscore in the right-most position of an unlabeled value which represents a blank. For example, 140\_-1409 is the range of all values starting with 140 in this particular variable.

## **Select Both Continuous Values and Groupings for a Search Line**

Since you cannot have both continuous values and groupings on the search line together, you must make two search lines for your search statement using the same variable name. To select on both continuous values and the defined groupings of a variable, use the following steps.

- 1 Make a single search line for either the continuous values or the groupings.
- 2 Choose the conjunction based on the operator you are using. If you use “is = to” then use “Or” for the conjunction. If you use “is not = to” then use “And” for the conjunction. These choices will ensure there is no conflict.
- 3 Write the second search line for the other type of selection.
- 4 Place a pair of parentheses around these two search lines by highlighting them and clicking on the ( .. ) button.

# Linking Search Lines to Make a Search Statement

Search statements may include multiple variables. Each selection line must be linked to the previous with a conjunction. Associated with multiple search lines is the use of Boolean conjunctions and parentheses. Parentheses define the order of operation and provide a larger degree of flexibility when building selection statements.

The process of using conjunctions and parenthesis for Boolean logic is optional. With each conjunction and group of parentheses, the complexity of the statement grows. To define selection statements with more than one variable, conjunctions are required and parentheses are highly recommended.

## **Precedence in a Search Statement**

The precedence of the Boolean operators is indicated by the order they are listed below. That is, parentheses have the highest precedence and Or has the lowest precedence. Precedence for multiple instances of the same operator is determined by the order it appears in the statement (first to last).

(. . )

Not

And

Or

# Conjunctions for Search Statements

A conjunction is used to link search lines to create a search statement. When adding a new search line to the statement, a conjunction is required. A conjunction may be selected by clicking on the And or Or buttons. Only one of the two Boolean conjunctions may be selected between lines.

## **Precedence in a Search Statement**

Versions 2.1.5 and earlier of SEER\*Stat did not support precedence between the And and Or Boolean operators. Without parentheses, these operators were evaluated in the order they appeared in the search statement. For example, consider this search statement:

Race = White Or Sex = Male And Year of Diagnosis=1980

SEER\*Stat versions 2.1.5 and earlier would yield a case subset of all whites (male and female) and all males (no matter the race) where they all have a diagnosis year of 1980.

To create precedence, you had to add parentheses to the search statement. For example, if you insert parentheses in the above search statement as shown:

Race = White Or (Sex = Male And Year of Diagnosis=1980)

then the case subset produced is all whites (male and female) regardless of their year of diagnosis and any males (no matter the race) with a diagnosis year of 1980. An entirely different subset from what was produced before.

Versions 3.0 and later support the Boolean operator precedence listed below.

(.. )

Not

And

Or

In SEER\*Stat version 3.0 and later, operators are evaluated in the order of their precedence rather than the order they appear in the search statement. Now, the first example from above:

Race = White Or Sex = Male And Year of Diagnosis = 1980

Would yield a case subset of all whites (male and female) regardless of their year of diagnosis and any males (no matter the race) with a diagnosis year of 1980. This is because And is evaluated before Or. Use parentheses to switch the precedence so that the Or is evaluated first.

That is, if you would like all whites (male and female) and all males (no matter the race) where the all have a diagnosis year of 1980, you have to add parentheses to the search statement to give the Or precedence.

(Race = White or Sex = Male) And Year of Diagnosis = 1980

## **Link Search Lines With a Conjunction (And/ Or)**

- 1 At the end of the first search line in the search statement, select a conjunction (And or Or).
- 2 Select a variable name, operator, and the variable groupings for the search line.
- 3 To change the conjunction, go to the end of the line you want to change and select the other conjunction.
- 4 To add more search lines, go to the end of the last line and select a conjunction. It may be necessary to press the Enter key on your keyboard to get to a new line.
- 5 When all of your new search lines have been added to the search statement, click OK.

## Add Parentheses to a Search Statement

Parenthesis around a group of search lines tells the computer to implement those search lines first when evaluating the search statement. By grouping search lines together and defining the order of operation, parentheses can provide a large degree of flexibility when building a search statement. Typically, it is easier to add parenthesis and define the order of operation in the search statement after all search lines have been added. To add parentheses to your search statement, follow these steps.

- 1 With your mouse, highlight the search lines which will be between the parentheses. Be careful if there is a Not in the statement. In some cases you will want the Not inside the parenthesis and in others you will want it outside the parenthesis.
- 2 Click on ( .. ), the Parenthesis button. Parenthesis will surround the search lines you had selected.

It is possible to edit the search statement like in a word processor through the Search Statements text box. Be extremely careful about editing the search statement in this way; there are many opportunities to make mistakes. However, there are times when this is the best way to work with parentheses. To change the location of parentheses in a search line follow these steps.

- 1 Highlight a parenthesis from the Search Statement box that must be moved.
- 2 Press the Delete key or Ctrl-X (Cut) on the keyboard to remove the parenthesis.
- 3 Place your cursor where the parenthesis should be and enter the parenthesis from the keyboard or use Ctrl-V (Paste) if you removed the parenthesis with Ctrl-X before.
- 4 Make sure all of your parentheses are paired appropriately.

## Add a “Not” to a Search Statement

Not tells the computer to select the subset which is the opposite of the group just described. For example the search statement:

Not (Race = White And Sex = Male)

will yield the subset of cases which are either non-white (male and female) or white females. The only people excluded are the white males, i.e., the group described inside the parentheses.

To add a Not to a search line use one of these methods.

- Place your cursor at the end of the search line and click the Not button. A Not should appear between the variable name and the last parenthesis.

- You may also use Cut (Ctrl-X), Copy (Ctrl-C), and Paste (Ctrl-V) to move/place a Not where you want it.

Typically, a Not is used with parentheses. If you need a negative for a single search line, you can use either the operator “is not = to” or a Not. Though either choice will work, the “is not = to” operator will be easier to handle if you need to add parentheses later. However, when creating a merged variable grouping, the Not button and the “is not = to” operator are disabled.

## Edit the Search Statement

You may need to edit a search statement that selects the subset of cases for your session or that defines a grouping of a merged variable. This is done on the Selection Window (or Edit Merged Grouping Window). Once on the Selection Window, make your edits and click OK.

- If you are selecting the subset of cases for your session on the Selection Tab, access the Selection Window by clicking the search statement’s Edit button.
- If you are editing a grouping for a merged variable, access the Edit Merged Grouping Window (Selection Window) from the Merged Variable Window by selecting the grouping you wish to edit and clicking the Edit button.

## Edit a Search Line or Change the Conjunction

- 1 In the search statement box, place your cursor anywhere in the search line you wish to edit. The current variable name, operator, and grouping selections for the line will be displayed in the Variable, Operator, and Selection(s) boxes.
- 2 Select the correct variable name, operator, and variable groupings from the boxes below.
- 3 Click the And or Or button. Your choice will replace the current conjunction.

## Add a Search Line

- 1 Place your cursor at the beginning or end of a search line. Press the Enter key on your keyboard to insert a blank line.
- 2 Move your cursor to the beginning of the blank line. If it is not the very first line of the search statement, select a conjunction (And or Or). If the new line is the first line, be sure to add a conjunction to the second line.

- 3 Select a variable name, operator, and the variable groupings for the search line.
- 4 To add another line, repeat steps 1-3.
- 5 When all of your new search lines have been added to the search statement, click OK.

## **Edit Parentheses in a Search Statement**

From the Selection Window (or Edit Merged Grouping Window) it is possible to edit the parentheses in a search line. To insert parenthesis in a search line follow these two steps.

- 1 With your mouse, highlight the search lines which will be between the parentheses. Be careful if a Not appears in your search statement. In some cases you will want the Not inside the parenthesis and in others you will want it outside the parenthesis.
- 2 Click the ( .. ) button. Parenthesis will surround the search lines you had selected.

It is possible to edit a search statement like in a word processor through the Search Statements text box. Be extremely careful about editing the search statement in this way; there are many opportunities to make mistakes. However, there are times when this is the best way to work with parentheses. To change the location of parentheses in a search line follow these steps.

- 1 Highlight a parenthesis from the Search Statement box that must be moved.
- 2 Press the Delete key or Ctrl-X (Cut) on the keyboard to remove the parenthesis.
- 3 Place your cursor where the parenthesis should be and enter the parenthesis from the keyboard or Ctrl-V (Paste).
- 4 Make sure all of your parentheses are paired appropriately.

## **Edit “Not” in the Search Statement**

Place your cursor at the end of the search line and click the Not button. A Not should appear between the variable name and the last parenthesis in your search statement. You may also use Cut (Ctrl-X), Copy (Ctrl-C), and Paste (Ctrl-V) to move/place a Not where you want it.

When creating or editing a merged variable grouping, the Not button is unavailable on the Merged Grouping Window (Selection Window).

## **Delete a Search Line from a Search Statement**

- 1 With the cursor, highlight the entire line you want to remove from your search statement.
- 2 Press the Delete key or Ctrl-X (Cut) to remove the line.

## **Order Search Lines in a Search Statement**

- 1 With the cursor, highlight the entire line you want to move in your search statement.
- 2 Remove the line by pressing Ctrl-X (Cut).
- 3 Place your cursor at the beginning or end of a search line. Press the Enter key on your keyboard to insert a blank line.
- 4 Move your cursor to the beginning of the blank line.
- 5 Paste to the line by pressing Ctrl-V (Paste).
- 6 Check that there is a conjunction at the beginning of each line but the first.

The Copy function (Ctrl-C) is also available for more complicated editing.

# Tips For Accurate Search Statements

The tips offered here should guide you in making accurate search statements. It is critical to carefully create search statements that capture the subset of cases you need for your analysis or merged variable grouping. In addition, inefficient statements can bog down the processing time.

- When defining the subset of cases for your session, a Selection Tab with no search statements selects all the cases in the database. Therefore, you should only make search statements for variables used to exclude cases.
- When creating a merged variable grouping, keep in mind that SEER\*Stat evaluates the search statement for the standard million and population files by setting each non-file variable to True.

## Select Variables Systematically

It is good practice to begin with the variables in the first category and work your way through the categories systematically. This will help you catch all of the conditions you want met. For example, in the SEER Public-Use database you might forget to consider the Behavior Code variable that indicates if a cancer is Malignant or In Situ.

Keep in mind that rate and trend calculations from the Frequency/Rate sessions require 3 pieces of data: cases, populations, and standard millions. Choose variables carefully to ensure that the numerators match the denominators and that the standard millions are for the appropriate cohort.

Standard millions are the age distributions for the standardizing population. When you subset your data based on age or display rates by age, you must use the age variable that is included in the standard million data file (Age recode).

Population counts are given by age, year, race, sex, and geographic area. Age is also in the standard millions and must be subset as described above. When you subset your data or display rates by year of diagnosis, race, sex, registry, or county you must use variables that are included in both the case and population files. In the SEER databases, the variable category called “Race, Sex, Year Dx, Registry, County” contains those variables.

There is a “Race and Age” category with variables that only appear in the case data. SEER collects cancer case data for single ages and for more races than are defined in the populations from the U.S Census Bureau. Do not use variables in the “Race and Age” category if you are calculating rates.



## Create an Exact Search Statement

You must be very thorough when specifying the cases you want for your analysis or merged variable. It is easy to inadvertently include cases which you do not intend to be a part of your subset. For example, if you are selecting a particular cancer from the SEER Public-Use database, consider whether it should have malignant behavior, in situ, or both.


For rate and trend calculations from the Frequency/Rate sessions it is especially important to consider if a cancer is sex specific to ensure the correct standard million and population subsets are used. It may seem redundant to select Site = Prostate and Sex = Male. If you do not select “sex = males,” the rates will be calculated as the count of Prostate cancer cases divided by the entire population (both males and females). This will half the rates. Instead make narrow search statements to collect only the cases, populations, and standard million subset that you need.

## Avoid Variable Groupings that Overlap

As you create your search statement, avoid choosing variable groupings that overlap. When you choose variable groupings that contain some of the same values, you are increasing the time for your job to run but not changing the results. Instead, you should choose variable groupings with mutually exclusive values. This becomes more of a issue when selecting a subset for a session than it is when creating a merged variable grouping.

For example, suppose you want statistics on people aged 45-74 and people age 65 and older. You create a user-defined variable which contains these two groupings. When writing your selection statement, the subset should not ask for both of these groups because they overlap. Instead, use the standard variable for age to select all people with ages greater than or equal to 45. Then use your user-defined variable to separate the groups on the Table Tab.

# Execute a Session

- 1 Click on the session window to activate it.
- 2 Click  or select Execute from the Session menu.

The time it takes to generate the requested statistics can range from a few seconds to several minutes. The Job Status Window displays the progress of the calculation process. While the job is executing, you can change the session or begin a new one without affecting the original job. It is possible to execute more than one job at a time.

## Job Status Window

When a session is executed, the Job Status Window appears. It contains the name of the session, the main statistic selected, the number of variables selected on the Table tab, and the database used for the calculation. In addition, two progress meters track the job through its numerous executable steps. The Current Operation progress meter names the current job step and reports the percent completed. The Overall progress meter indicates what percent of the overall job steps have been completed. Once the job is complete, the Job Status Window will be replaced by the output matrix.

It is easy to misinterpret the Overall progress meter, for it is not based on time as one might assume. Instead, the Overall progress meter reports the percent of job steps completed, not the time any one of the jobs will take. For example, if the Overall progress meter indicates that 50% of the job steps have been completed 10 seconds after you executed your session, this does not mean that there is only 10 seconds left in the job. Actually there could be either more or less time left. You cannot predict execution time with the Overall progress meter.

If you are utilizing SEER\*Stat within the client-server configuration, the Finish Remote button can be used on jobs that take more than several minutes to complete. Once the execution is complete, an e-mail message will be sent notifying you of the location of your resultant matrix.


## Finish Remote

The Finish Remote button is only visible on the Job Status Window if you are utilizing SEER\*Stat within the client-server configuration.

The finish remote functionality is generally used for analyses that take more than several minutes to complete. By clicking the Finish Remote button, you are requesting that the execution continue on the server but you do not want status updates in Job Status Window. With the Job Status Window no longer present, you can exit SEER\*Stat and even turn

off your PC while the server continues the execution. Once the execution completes on the server, an email message will be sent notifying you of the location of your resultant matrix. Matrices are available on a password protected Web site, and the specific URL or address will be provided in the email notification. Your username and password, needed to download the matrices with your Web browser, are the same as your SEER\*Stat username and password.

## Cancel a Job

- 1 Click  on the tool bar, select Cancel from the Job menu, or click the Cancel button at bottom of Job Status Window.
- 2 Click Close on the Job Status Window.

## Set the Job Priority

When more than one job is running, you can control the amount of CPU time allotted to the jobs by setting the Job Priority. There are three levels: High, Normal, and Low.



- You may set the priority of a job while it is executing through the Job menu.
- You may change the default job priority in Preferences from the File menu.

# Output Matrix


The output matrix displays the statistics/values you requested in your session. When a session finishes executing, the associated output matrix is automatically displayed. You can change the appearance of the output matrix, save it, print it, copy it to the Windows clipboard and/or export the statistics/values so that they may be used in another application. When an output matrix is saved, the associated session is saved with it. This is particularly useful because it limits the number of files on the computer and helps you regenerate the session when necessary. The session may be retrieved from the output matrix, modified if necessary and executed again.

## Open a Matrix

To open an output matrix, you may do any of the following.

- Click  to open a previously saved matrix. The file extensions for a Frequency/Rate Matrix, Survival Matrix, and Case Listing Matrix respectively are '.sim', '.ssm', and '.slm'.
- While on a Frequency/Rate Session, Survival Session, or a Case Listing Session, click  to execute the session and create the matrix.
- Use the File menu command Open.

## Save a Matrix

- 1 Click on the output matrix to activate it.
- 2 Click  or use the Save or Save As commands on the File menu.

A matrix file may be saved at any time. A saved matrix can be retrieved and modified at a later time. It is not necessary to keep both the session and the matrix. When the output matrix is saved, the associated session is saved with it and can be recalled, or sessions may be executed again to recreate the matrix.

## View the Session Which Produced the Matrix

When the output matrix is created, the session which produced it is linked to the matrix. This session can be recalled using the View Session command from the Matrix menu. However, as with any newly opened session, database-wide variables are updated to the current definition. This could cause a different output matrix to be created from the session. If the session is executed, the output matrix will not reflect any filtering or formatting changes that were made to the original matrix.

# Matrix Properties Window

Through the Matrix Properties, SEER\*Stat provides details about the output matrix. The Matrix Properties may be accessed by choosing Properties from the Matrix menu.

## **Database**

Lists the database used to create the matrix. This is the same database that was selected on the Data Tab of your session.

## **Statistic**

Lists the statistic displayed in the matrix that was selected on the Data Tab of your Frequency/Rate or Survival session, or indicates that the matrix is a Case Listing.

## **File Name**

Once the matrix has been saved, its file name appears in this field.

## **Directory**

Once the matrix has been saved, the full directory path for the matrix file appears in this field.

## **File Size**

Once the matrix has been saved, the size of the matrix file in bytes appears in this field.

## **Cell/Page/Row Count**

For a Frequency/Rate matrix the number of cells making up your matrix is given here. The number of pages is reported for a Survival matrix and the number of rows for a Case Listing matrix.

## **Creation Time**

This is the time in seconds it took to create the matrix.

## **Created**

This is the date when the matrix was first saved.

## **Modified**

This is the date when the matrix was last saved.

## **Accessed**

On the electronic copy of the Matrix Properties Window, this will always be today's date, since you must have accessed the matrix to look at its properties. However, when printed, this is much like the print date.

## **Information box**

If you calculated frequencies/rates, the Information box displays the row, column, and page variables. If you calculated survival statistics, the

Information box displays the page variables and all of the exclusion counts. And, if you executed a Case Listing session, column and sort variables are listed here.

## Frequency/Rate Matrix

### Frequency/Rate Matrix Window

The output matrix window is made up of the title, row and column headings, the statistics, and the footnotes. In addition, the Matrix menu and Right Mouse Button menu are available when the matrix window is active. Here is a short description of each feature.

#### **Title**

The title displayed in a matrix window is the first line of the matrix title from the Output Tab. You may change the title on the Matrix Options window. If you change the title in Matrix Options and then recall the session, the session contains the original, not the edited title.

#### **Row and Column Headings**

The row and column headings describe which statistics are being displayed. It is possible to change the order of the row or column variables from the Order Labels window. You may also select the columns of statistics shown from the Matrix Options window.

#### **Statistics**

The statistics are the actual output from SEER\*Stat. You can change the appearance of the numbers by using the Filter.

#### **Footnotes**

Footnotes and flags are added to the matrix to explain calculation procedures, blank cells, options selected in the session, and tests run on the data results. Footnotes and flags may be removed on the Matrix Options window.

#### **Matrix Menu and Right Mouse Button Menu**

The Matrix menu commands allow you to do quite a bit with your matrix. Any command with RMB next to it is also available from the Right Mouse Button Menu. To access the Right Mouse Button Menu, highlight a column or row and click on it with the right mouse button. The Right Mouse Button Menu will appear.

#### **View Session**

View the session that created the matrix.

#### **Options**


Access the Matrix Options window.

#### **Font**

Change the matrix font.

<b>Order (RMB)</b>	Change the order of the row or column variables.
<b>Filter (RMB)</b>	Access the Filter window to change the appearance of the matrix numbers (in columns only).
<b>Lock (RMB)</b>	Lock a row and column in place for easy comparisons within the matrix window.
<b>Page</b>	Turn pages in the matrix.
<b>Export</b>	Export the matrix as a text file for use in other applications.
<b>Properties</b>	View the Matrix Properties window.

## Frequency/Rate Matrix Options

To access the Matrix Options window, click on  or select Options from the Matrix menu.

### **Titles**

If you wish to add a title or edit the title of your output matrix, make your changes in the text box provided on the Options Window. Only the first line of the title will be displayed on the SEER\*Stat matrix window. To see how the matrix title will print, select Print Preview from the File menu. When the matrix is printed, the text will be centered on the page, but will not wrap to a second line without carriage returns. To get a multi-lined title, press Enter at the end of each line of text.

### **Statistics**

Select the statistics you wish to appear on the matrix. Only statistics available on your matrix will be active. Notice that you cannot display the standard errors without displaying the confidence intervals.

### **Display Titles**

This option removes or reveals the titles in the matrix. Titles may be edited at the top of the window.

### **Display Flags**

This option removes or reveals the flags in the matrix. Each flag is related to a footnote.

### **Display Footnotes**

This option removes or reveals the footnotes. Footnotes are included in the matrix to explain calculation procedures, blank cells, options selected in the session, and tests run on the data results. Footnotes cannot be displayed without the flags to which they refer.

### **Hide Zero Count Rows**


When selected, the rows with zeroes in the count (not population) will be removed from the matrix display. A row is hidden only if all counts for that row in all pages are zero.

### **Hide PC Rows, Hide EAPC Rows, Hide Rates**

When trends have been calculated, these options remove or reveal the percent change, the estimated annual percent change, and the rates calculated to determine the trends respectively. You cannot hide all three row types at once. At least one row type must be shown. If you have trend matrices from versions of SEER\*Stat prior to 2.1, the Hide PC Rows and Hide EAPC Rows options will not work.

## **Matrix Footnotes**

In a Frequency/Rate or Survival matrix, footnotes and flags are added to the output matrix to explain calculation procedures, blank cells, options selected in the session, and tests run on the data results. To remove the footnotes:

- 1 Select Options from the Matrix menu or click on  .
- 2 Click the Hide Footnotes option.

## **Frequency/Rate Matrix Navigation**

To manipulate the output matrix, use the directions found in the following sections.

Page through a Matrix

Select Columns or Rows in a Matrix

Lock Feature for a Matrix

Resize Columns or Rows in a Matrix

Reorder Column or Row Variables in a Matrix

## **Page Through a Matrix**

Only one page of the output matrix is displayed in the window at a time. The horizontal and vertical scroll bars allow you to move through the page.

To change the matrix page that is displayed, click the drop-down list box on the toolbar that displays the name of the current matrix page. From this box, select the name of the page that you wish to view.

You can also use the database-style buttons on the toolbar to switch matrix pages. The buttons allow you to jump to the first, previous, next and last page of the matrix. In addition, the paging controls can be found



under Page in the Matrix menu. The paging controls are disabled if the report is a one-page report.

## Select Columns or Rows in a Matrix

You can select columns or rows in the output matrix for an operation by highlighting them. For example, to filter a column, you must first highlight it. Here is a summary of the methods used to highlight columns and rows.

- To select a single column or row, click the left-most or top-most label.
- To select adjacent columns or rows drag across the column or row labels. Alternately, select the first column or row; then hold down the Shift key and select the last column or row label.
- To select non-adjacent columns or rows select the first column or row by clicking on the left-most or top-most label. Then hold down the Ctrl key while you select the other columns or rows.
- To select all columns or rows choose Highlight All from the Edit menu and select either Columns or Rows.

## Lock Feature for a Matrix

SEER\*Stat allows you to lock both columns and rows in a Frequency/Rate or Survival Matrix. Locking a column means that the column and all the columns to the left of it will not scroll off the left side of the window, but instead are constantly in view. In the same way, locking a row will hold the row and all rows above it so that they cannot be scrolled off the top of the window. To lock columns or rows:

- 1 Select the columns or rows you wish to lock.
- 2 Choose Lock from the Matrix menu or Right Mouse Button menu.

A thick vertical or horizontal bar is displayed to indicate that the columns or rows are locked. To unlock previously locked columns and rows, highlight the locked columns or rows and select Lock again from the Right Mouse Button menu. This turns off the lock.

Locking columns or rows can be very useful. For example, if you have totals in the first column or row of your table and there are too many columns or rows to all fit in the window, locking the totals in place could be helpful. If you use the scroll bars to view the later columns or rows in the matrix, the total information scrolls off the window. By locking the total, you can always see the total information.

## Resize Columns or Rows in a Matrix

You may resize the rows or columns of the output matrix to any height or width that you want. To resize a row or column, place the cursor over the boundary lines for the row or column labels. When you are in the right

spot the cursor changes its shape to a line with arrows coming out of each side. This is called the resizing cursor. Once the cursor is in this shape click and drag the row or column to the desired size.

You may resize multiple rows or columns at once. To do this, select the rows or columns you wish to resize. Then resize any one of them as described above. All of the selected rows or columns will be resized to the same size.

## **Reorder Column or Row Variables in a Frequency/Rate Matrix**

- 1 On a Frequency/Rate output matrix, select any column or row.
- 2 Choose Order from the Matrix menu or Right Mouse Button menu.

The Order Labels Window will appear.

## **Order Labels Window**

Select the variable label you wish to move on the output matrix. Press the Move Up and Move Down buttons until the variable label is where it should be. Click OK when you are finished.

## **Survival Matrix**

### **Survival Matrix Window**

The output matrix window is made up of the title, row and column headings, the statistics, and the footnotes. In addition, the Matrix menu and Right Mouse Button menu are available when the matrix window is active. Here is a short description of each feature.

#### **Title**

The title displayed in matrix window is the first line of the matrix title from the Output Tab. You may change the title on the Matrix Options window. If you change the title in Matrix Options and then recall the session, the session contains the original, not the edited title.

#### **Row and Column Headings**

The row and column headings describe which statistics are being displayed. It is possible to select the columns of statistics shown from the Matrix Options window.

#### **Statistics**

The statistics are the actual output from SEER\*Stat. You can change the appearance of the numbers by using the Filter.

## Footnotes


Footnotes and flags are added to the matrix to explain calculation procedures, blank cells, options selected in the session, and tests run on the data results. Footnotes and flags may be removed on the Matrix Options window.

## Matrix Menu and Right Mouse Button Menu

The Matrix menu commands allow you to do quite a bit with your matrix. Any command with RMB next to it is also available from the Right Mouse Button Menu. To access the Right Mouse Button Menu, highlight a column or row and click on it with the right mouse button. The Right Mouse Button Menu will appear.

<b>View Session</b>	View the session that created the matrix.
<b>Options</b>	Access the Matrix Options window.
<b>Font</b>	Change the matrix font.
<b>Order (RMB)</b>	Only available in a Frequency/Rate matrix.
<b>Filter (RMB)</b>	Access the Filter window to change the appearance of the matrix numbers (in columns only).
<b>Lock (RMB)</b>	Lock a row and column in place for easy comparisons within the matrix window.
<b>Page</b>	Turn pages in the matrix.
<b>Export</b>	Export the matrix as a text file for use in other applications.
<b>Properties</b>	View the Matrix Properties window.

## Survival Matrix Options

To access the Matrix Options window, click on  or select Options from the Matrix menu.

### Titles

If you wish to add a title or edit the title of your output matrix, make your changes in the text box provided on the Options Window. Only the first line of the title will be displayed on the SEER\*Stat matrix window. To see how the matrix title will print, select Print Preview from the File menu. When the matrix is printed the text will be centered on the page, but will not wrap to a second line without carriage returns. To get a multi-lined title, press Enter at the end of each line of text.

### Statistics

Select the statistics you wish to appear in the matrix. Only statistics available to your matrix will be active.

**Display Titles**

This option removes or reveals the titles in the matrix. Titles may be edited at the top of the window.

**Display Flags**


This option removes or reveals the flags in the matrix. Each flag is related to a footnote.

**Display Footnotes**

This option removes or reveals the footnotes. Footnotes are included in the matrix to explain calculation procedures, blank cells, options selected in the session, and tests run on the data results. Footnotes cannot be displayed without the flags to which they refer.

## Matrix Footnotes

In a Frequency/Rate or Survival matrix, footnotes and flags are added to the output matrix to explain calculation procedures, blank cells, options selected in the session, and tests run on the data results. To remove the footnotes:

- 1 Select Options from the Matrix menu or click on .
- 2 Click the Hide Footnotes option.

## Survival Matrix Navigation

To manipulate the output matrix, use the directions found in the following sections.

Page through a Matrix

Select Columns or Rows in a Matrix

Lock Feature for a Matrix

Resize Columns or Rows in a Matrix

## Page Through a Matrix

Only one page of the output matrix is displayed in the window at a time. The horizontal and vertical scroll bars allow you to move through the page.

To change the matrix page that is displayed, click the drop-down list box on the toolbar that displays the name of the current matrix page. From this box, select the name of the page that you wish to view.

You can also use the database-style buttons on the toolbar to switch matrix pages. The buttons allow you to jump to the first, previous, next and last page of the matrix. In addition, the paging controls can be found

under Page in the Matrix menu. The paging controls are disabled if the report is a one-page report.

## Select Columns or Rows in a Matrix

You can select columns or rows in the output matrix for an operation by highlighting them. For example, to filter a column, you must first highlight it. Here is a summary of the methods used to highlight columns and rows.

- To select a single column or row, click the left-most or top-most label.
- To select adjacent columns or rows drag across the column or row labels. Alternately, select the first column or row; then hold down the Shift key and select the last column or row label.
- To select non-adjacent columns or rows select the first column or row by clicking on the left-most or top-most label. Then hold down the Ctrl key while you select the other columns or rows.
- To select all columns or rows choose Highlight All from the Edit menu and select either Columns or Rows.

## Lock Feature for a Matrix

SEER\*Stat allows you to lock both columns and rows in a Frequency/Rate or Survival Matrix. Locking a column means that the column and all the columns to the left of it will not scroll off the left side of the window, but instead are constantly in view. In the same way, locking a row will hold the row and all rows above it so that they cannot be scrolled off the top of the window. To lock columns or rows:

- 1 Select the columns or rows you wish to lock.
- 2 Choose Lock from the Matrix menu or Right Mouse Button menu.

A thick vertical or horizontal bar is displayed to indicate that the columns or rows are locked. To unlock previously locked columns and rows, highlight the locked columns or rows and select **Lock** again from the Right Mouse Button menu. This turns off the lock.

Locking columns or rows can be very useful. For example, if you have totals in the first column or row of your table and there are too many columns or rows to all fit in the window, locking the totals in place could be helpful. If you use the scroll bars to view the later columns or rows in the matrix, the total information scrolls off the window. By locking the total, you can always see the total information.

## Resize Columns or Rows in a Matrix

You may resize the rows or columns of the output matrix to any height or width that you want. To resize a row or column, place the cursor over the boundary lines for the row or column labels. When you are in the right

spot the cursor changes its shape to a line with arrows coming out of each side. This is called the resizing cursor. Once the cursor is in this shape click and drag the row or column to the desired size.

You may resize multiple rows or columns at once. To do this, select the rows or columns you wish to resize. Then resize any one of them as described above. All of the selected rows or columns will be resized to the same size.

## Case Listing Matrix

### Case Listing Matrix Window

The output matrix window is made up of the title, row and column headings, and the variable values. In addition, the Matrix menu and Right Mouse Button menu are available when the matrix window is active. Here is a short description of each feature.

#### **Title**

The title displayed in the matrix window is the first line of the matrix title from the Output Tab. You may change the title on the Matrix Options window. If you change the title in Matrix Options and then recall the session, the session contains the original, not the edited title.

#### **Row and Column Headings**

The row headings show the record numbers in the Case Listing matrix and the column headings are the chosen column variables. It is possible to change the order of the column variables in the Matrix Options window. You may also choose to hide columns using the Right Mouse Button menu or the Matrix Options window.

#### **Variable Values**

The variable values are the actual output for a Case Listing in SEER\*Stat. You can change the appearance of the values using the Display As option from the Matrix menu.

#### **Matrix Menu and Right Mouse Button Menu**

The Matrix menu commands allow you to do quite a bit with your matrix. Any command with RMB next to it is also available from the Right Mouse Button Menu. To access the Right Mouse Button Menu, highlight a column or row and click on it with the right mouse button. The Right Mouse Button Menu will appear.

#### **View Session**

View the session that created the matrix.

#### **Options**


Access the Matrix Options window.

#### **Font**

Change the matrix font.

<b>Copy</b>	Copies selected variable values to the Windows Clipboard
<b>Display Row Numbers (RMB)</b>	Displays or hides row numbers.
<b>Alignment (RMB)</b>	Changes the alignment of the variable values within the column.
<b>Display As (RMB)</b>	Changes the appearance of the values in columns to Formatted, Unformatted, or Converted.
<b>Auto-size (RMB)</b>	Makes the column width fit the contents.
<b>Lock (RMB)</b>	Lock a column in place for easy comparisons within the matrix window.
<b>Hide (RMB)</b>	Hides columns in the matrix.
<b>Export</b>	Export the matrix as a text file for use in other applications.
<b>Properties</b>	View the Matrix Properties window.

## Case Listing Matrix Options



To access the Matrix Options Window, click on  or select Options from the Matrix menu.

### Title

If you wish to add or edit the title of your output matrix, make your changes in the text box provided on the Options Window. Only the first line of the title will be displayed on the SEER\*Stat matrix window. To see how the matrix title will print, select Print Preview from the File menu. When the matrix is printed, the title will be centered on the page. To get a multi-lined title, press Enter at the end of each line of text.

### Hidden and Displayed Variables

The Hidden and Displayed Variables boxes allow you to hide or rearrange columns in your output matrix.

To hide column(s), select the variable(s) you wish to hide from the list of Displayed variables and use the arrow button, , between the text boxes to move them to the Hidden box. To move all variables from one text box to the other, click the double arrows button, , in the direction you want the variables to move. You may also hide columns using the Hide option on the Right Mouse Button menu.

To rearrange the order of columns in the output matrix, select the variable to be moved from the Displayed Variables box and click the Up or Dn buttons until the variable has reached the correct position.

# Case Listing Matrix Navigation

To manipulate the output matrix, use the directions found in the following sections.

Select Columns or Rows in a Matrix

Change Alignment of Values in a Matrix

Lock Feature for a Matrix

Hide Columns in a Matrix

Change Format of Variable Values

Resize Columns or Rows in a Matrix

Reorder Column Variables in a Matrix

## Select Columns or Rows in a Matrix

You can select columns or rows in the output matrix for an operation by highlighting them. For example, to filter a column, you must first highlight it. Here is a summary of the methods used to highlight columns and rows.

- To select a single column or row, click the left-most or top-most label.
- To select adjacent columns or rows drag across the column or row labels. Alternately, select the first column or row; then hold down the Shift key and select the last column or row label.
- To select non-adjacent columns or rows select the first column or row by clicking on the left-most or top-most label. Then hold down the Ctrl key while you select the other columns or rows.
- To select all columns or rows choose Highlight All from the Edit menu and select either Columns or Rows.

## Change Alignment of Values in a Matrix

Variable values in a Case Listing matrix have a default alignment. Formatted and Converted values are left-aligned, and Unformatted values are right aligned. Changing the alignment does not change the format of the values. To change the alignment of variable values in the matrix:

- 1 Select the column(s) to realign.
- 2 Choose Alignment from the Matrix menu or the Right Mouse Button menu.
- 3 Select Left, Right, or Center to change the alignment in the selected column(s).



## Lock Feature for a Case Listing Matrix

SEER\*Stat allows you to lock columns in a Case Listing output matrix. Locking columns means that all columns you select (adjacent or non-adjacent) will be moved to the left of the matrix and will not scroll off the matrix window. Those selected columns are constantly in view. To lock a column(s):

- 1 Select the columns you wish to lock.
- 2 Choose Lock from the Matrix menu or the Right Mouse Button menu.

A thick vertical bar indicates that all columns to the left of it are locked. To unlock previously locked columns, highlight the locked column that borders the thick vertical line and select Lock again from the Right Mouse Button menu. The unlocked columns do not return to their original order.

### Note



You cannot lock rows in a Case Listing matrix.

## Hide Columns in a Case Listing Matrix

There are two ways to hide columns in a Case Listing Matrix. First using the menus:

- 1 Select the column or columns you wish to hide.
- 2 Choose Hide from the Matrix menu or the Right Mouse Button menu .
- 3 To display the hidden columns, choose Show from the Matrix or Right Mouse Button menu and select the column(s) you want to reveal.

Columns may also be hidden using the Case Listing Matrix Options Window.

- 1 Open the Matrix Options Window by clicking on  or selecting Options from the Matrix menu.
- 2 Select the column variable(s) you wish to hide from the Displayed Variables box.
- 3 Move the variables to the Hidden Variables box using the arrow button,  , between the two variable boxes.
- 4 To display the Hidden variables, use the arrow button to move the selected variables back to the Displayed Variables box.

## Change Format of Variable Values in a Matrix

Variable values in a Case Listing matrix are output as formatted values by default. To change the format of variable values:

- 1 Select the column(s) in which you wish to change the values.

- 2 Choose Display As from the Matrix menu or the Right Mouse Button menu.
- 3 Select whether you want to display the values as Formatted, Unformatted, or Converted.

When you change the format of the variable values, the columns are auto-sized to fit the cell contents and the alignment is changed to the default display type.

**Note**

Only continuous or conversion variables (i.e. Histology, Year of Birth) may be displayed as Converted. This option is disabled when other column variables are selected.

## Resize Columns or Rows in a Case Listing Matrix


You may resize the rows or columns of the Case Listing matrix to any height or width that you want. To resize a column, place the cursor over the boundary line for the column labels. When you are in the right spot the cursor changes its shape to a line with arrows coming out of each side. This is called the resizing cursor. Once the cursor is in this shape click and drag the column to the desired size. You can use the Auto-size option from the Matrix menu or the Right Mouse Button menu to make the column width automatically fit the size of the cell contents.

You may resize multiple columns at once. To do this, select the columns you wish to resize. Then resize one of them as described above. All of the selected columns will be resized to the same size.

You can resize rows using the same procedures for columns, but changing the height of one row applies the new height to all rows in your matrix.

## Reorder Column Variables in a Case Listing Matrix

To reorder the columns in a Case Listing matrix:

- 1 Open the Matrix Options window by clicking on  or selecting Options from the Matrix menu.
- 2 Select the column variable(s) to be moved within the Displayed Variables box and click the Up or Dn buttons until the variable(s) reaches the correct position.

## Change Matrix Font

SEER\*Stat uses the standard Windows font dialog to change the font, font size, and style for the matrix. To access the Font Window, select Font

from the Matrix menu. You may change the default font on the Preferences Window.

## Font Window

The Font Window allows you to change the font type, font size, and font style for your output matrix.

### Font Type

The Font box lists all the fonts that are available. The available fonts depend on the installed TrueType fonts as well as resident fonts within the printer you have selected. When the box first appears, the font already in use for the matrix is highlighted. Click the font you wish to use (if different from the highlighted font).

### Font Style

The Font Style box lists the styles that you can assign to the font selected. To change the style, click on the style you wish to use. Your choices include some or all of the following.

Regular

**Bold**

*Italic*

***Bold Italic***

### Font Size

The Font Size box lists common point sizes for the highlighted font. When the box first appears, the point size for the font already in use for the selected object is highlighted, and the highlighted point size appears in the edit box at the top. Click the point size you want (if different from the highlighted size). You can select directly from the list or type the new point size in the edit box at the top (if you know that additional sizes are available).

### Font Sample

The Sample box displays a sample of the font you have selected. The sample shows the font, style, and size you have specified. You can use this box to preview the results as you experiment with different formatting options.

# Number Filter

In a Frequency/Rate or Survival Matrix, the filter is used to change the way that numbers are displayed in the matrix. As the name indicates, filtering data is a way of adding characters to a field at a specified place or of removing unwanted characters from a field for display purposes only. It is important to note that applying a filter does not change the underlying data, only the way in which the data is displayed.

## Accessing the Number Filter

While viewing the matrix, you may filter any column. To specify a filter for a column, follow these steps.

- 1 Select the column(s) that you want to filter.
- 2 Select Filter from the Matrix menu or the Right Mouse Button menu to access the Filter Window.

### Note

The Filter option on the Matrix menu or the Right Mouse Button menu is only available for Frequency/Rate or Survival matrices.

## Filter Window

The filter is used to change the way that numbers are displayed in the Frequency/Rate or Survival matrix. On the Filter Window the following options are available.

### Suppress if Zero

This option displays a blank if the cell value is zero.

### Thousands Separator

The thousands separator option allows you to choose whether you want a number over 999 to be separated by commas.

### Leading Zero

Leading Zero automatically inserts a zero to the left of the decimal point in decimal amounts which fall between negative one and one.

### Fixed Width

When Fixed Width is checked, SEER\*Stat sets the width for certain numbers in your matrix, aligns numbers to the right or to the left, and uses the supplied characters to pad numbers which are smaller than the width.

### Width

This represents the width of the number (not the cell) which is displayed.

**Fill Side**

The fill side option lets you specify which side to fill with the user specified characters.

**Fill Character**

The fill character option lets you enter one character which fills any remaining space up to the width selected. The default fill character is a space. To change the fill character, delete the space and enter the new character.

**Rounding**

Rounding allows you to round numbers to a specific decimal place. You can round on either side of the decimal point.

**Negatives**

The Negatives option allows you to choose the way you want negative numbers to appear. You have three options in how negatives are displayed: a leading minus sign, a trailing minus sign, or parentheses to indicate a negative.

**Prefix/ Suffix Characters**

When checked, every number is displayed with the supplied prefix and/or suffix character(s).

**Sample**

Shows a number -5,555,555.55 displayed using the current filter settings.

## **Suppress if Zero Filter**

This filter option displays a blank if the cell value is zero. This is useful when your matrix contains a large amount of zero values, and you only want to see the non-zero elements.

## **Thousands Separator Filter**

The thousands separator filter option allows you to choose whether you want a number over 999 to be separated by commas. When this option is checked, the comma ( , ) is inserted as the thousands separator character for amounts over 999. Your numbers are displayed as:

1,000

10,000

999,000

When unchecked, the thousands separator character is not used. Your numbers would then be displayed as:

1000

10000

999000

## Leading Zero Filter

Leading Zero filter automatically inserts a zero to the left of the decimal point in decimal amounts which fall between -1 and 1, non-inclusive. With this option checked, a decimal amount which falls in this range is displayed with a leading zero:

0.001

0.99999

0.755

When this option is unchecked, decimal amounts within the range are displayed without a leading zero

.001

.99999

.755

## Fixed Width Filter

When the filter option, Fixed Width, is checked, SEER\*Stat sets the width for certain numbers in your matrix and uses the supplied characters to pad numbers which are smaller than the width. This has the effect of not only aligning numbers to the right or to the left, but also of giving every number the same width. The number is filled with asterisks (\*) if the number takes up more positions than the designated width.

This option should be used when you want to fill the empty spaces with something other than a blank or as a quick way of identifying numbers which extend outside a particular range. Fixed width is also useful if you have specified a prefix and want the numbers to be aligned to the right but want the prefix to be aligned to the left.

### Width

This represents the width of the number (not the cell) which is displayed. It can be any width from 1 to 32. For example, if you select 4, then every value of that number type is displayed with a width of 4. If an underlying number has a width of greater than 4, asterisks are displayed. Prefix and suffix characters are NOT counted when determining whether a number fits into a fixed width.

### Fill Side

The fill side option lets you specify which side to fill with the user specified characters. You can fill or pad to the left or to the right of the number.

### **Fill Character**

The fill character option lets you enter one character which fills any remaining space up to the width selected. For example, a fixed width number with a width of 10, left fill side, and using “0” as the fill character would display “55555” as “0000055555”. The default fill character is a space. To change the fill character, delete the space and enter the character you want.

## **Rounding Filter**

The Rounding filter allows you to round numbers to a specific decimal place. You can round on either side of the decimal point. All selections, which are greater than one, round to the selected place value on the left of the decimal point. All selections, which are less than one, round to the selected place value on the right of the decimal point. For example, if you select 0.01 and your number is 555.555, the number displayed would be 555.56. Similarly, if you select 100 with the same number, the displayed number would be 600.

## **Negatives Filter**

The Negatives filter option allows you to choose the way you want negative numbers to appear. You have three options in how negatives are displayed:

-123 displays a leading minus sign

123- displays a trailing minus sign

(123) displays parentheses to indicate a negative

## **Prefix Characters Filter**

When this filter option is checked, every number is displayed with the supplied prefix character. No space is inserted between the specified prefix character and the number to be displayed. A value of 32.45 with a prefix character of ‘\$’ is displayed as \$32.45. If fixed width is specified, the prefix character does not contribute to the width of the number.

## **Suffix Characters Filter**

When this filter option is checked, every number is displayed with the supplied suffix character. As with the prefix, no space is inserted between the number and the suffix character. Since more than one suffix character can be specified, you can use one of them to supply a space if one is desired. For example, a value of 98 with suffix characters of ‘ %’ would be displayed as 98%. Suffix characters do not contribute to the width of a number if fixed width is specified.

# Sharing Your Results with Other Users and Applications

## Copy a Matrix to the Windows Clipboard

### Frequency/Rate and Survival Matrices

You can copy a cell, a single page, or all pages from a Frequency/Rate or Survival matrix to the Windows clipboard by selecting Copy from the Edit menu or by pressing Cntrl+C for a cell, Cntrl+P for a page, or Cntrl+A for all pages.


### Case Listing Matrix

You can copy a single row, column, or multiple rows and columns from a Case Listing Matrix to the Windows clipboard. Select individual or nonadjacent rows and columns in the matrix window or select Highlight from the Edit menu to highlight all rows or columns in the matrix. Use the Copy option from the Edit menu or the Right Mouse Button menu to copy the selected data.

The Copy feature can be very helpful to quickly paste SEER\*Stat data into other Windows applications, such as word processors and spreadsheets. The data will be tab delimited, that is, the data columns from the matrix will be separated not by lines, but by tabs.

## Print a Session or Matrix

You may print hard copies of both sessions and matrices. A session print out lists the database, statistics, selections, table variables, user-defined variable definitions, and options you made in the session. Matrix pages are printed exactly as shown on screen, including any formatting options and font selection.

- Select Print Preview from the File menu to see what you will be printing.
- Select Print from the File menu, or click  on the toolbar to print the contents of the currently active session or matrix.
- Select Print Setup from the File menu to set up the printing parameters.

### Note

If your matrices are printing too slowly, try selecting a font which is resident in your printer, or hide statistics/variables you do not need printed



using the Matrix Options. In addition, try printing without the Use Shading When Printing option from the Preferences dialog.

## Use Shading When Printing

The shading when printing option of the Preferences dialog adds shading to the column and row titles of the printed output matrix. If you plan to make copies or FAX the matrix, it may be wise to turn off this option.

## Print User-Defined Variable Definitions

The check box labeled “Print User-defined Variable Definitions” on the Preferences dialog allows you to choose whether or not to print the definitions of all user-defined variables used in a session. If selected, the title of the user-defined variable, the standard variable on which it is based, the groupings and all values associated with each grouping are listed on the session print out. This option may be selected or deselected before or after a job has been executed. Turn this option off to reduce the length of your printout. Turn it on in order to share your variable definitions with other users or to review the definitions of your variables.

### Note

If you have turned off the print user-defined variable definitions option, SEER\*Stat will still list the title of the user-defined variable and standard variable on which it is based in the Session print out.

## Exporting Results to Other Applications

To Export Data:

- 1 Click on the matrix to activate it.
- 2 Select Export from the Matrix menu. The Export Window will appear.
- 3 On the Export Window, select the destination and options which will produce the most compatible ASCII text file for the software application you plan to use.

During export, all other SEER\*Stat processes are stopped. For very large matrices, you may choose to minimize the SEER\*Stat Window by clicking the minimize button of the Export Window. The SEER\*Stat button on the Windows taskbar will indicate the progress of the export. SEER\*Stat will not automatically maximize when the export process is complete.

## Export Data Files

SEER\*Stat gives you the option to create a text-based data file containing all the statistics/values generated in your analysis. The exported data file can be used as input into other standard software packages.

When you export a matrix, two output files are created.

- The output text file with extension “.txt” which contains all the statistics/values requested and the table variables in a tab delimited format
- A text-based data dictionary with extension “.dic” consisting of the table values

Options on the Export WindowExportWindow>proc include the following.

- The name and path for both the exported text file (.txt) and the text-based data dictionary file (.dic)

The exported text file may be saved as a gzipped file if the file extension “.gz” is placed at the end of the exported text file name.

- The representation of output values as numbers, labels, or labels with quotes (Numeric Representation option is disabled in a Case Listing)
- The field and line delimiters (separators)
- The representation of missing characters as spaces or periods (Missing Characters option is disabled in a Case Listing)
- To enclose fields containing delimiter in quotes
- To remove all thousands separators (commas)
- To remove flags and footnote characters (this option is disabled in a Case Listing)
- To output variable names before data
- To set the options as default or return to the default settings

## Export Data Dictionary

The output data dictionary gives information about the exported data text file. There are four sections of the output data dictionary: the filenames, the variables and statistic/values, export options, and the variable recodes.

### **Filenames**

Lists the path name for the text file and the path name for the original matrix.

### **Export Options**

Lists the options you selected on the Export Window. These details about the exported data file include:

- Whether the data file produced is gzipped (.gz)
- Variable format: numeric, labels, or labels with quotes
- Line delimiter or File format: Dos/Windows or UNIX
- Field delimiter: tab, space, comma, or semi-colon
- Missing character: space or period
- Whether fields containing the delimiter are enclosed in quotes
- Whether the thousands separators are removed
- Whether the flags for the footnotes were removed
- Whether variable names are output before data

### **Variables and Statistics/Values**

Lists the variables names and statistics/values included in the text file in the order of appearance from left to right. If a variable is user-defined, the base variable is also listed. The order of the table variables is standard for all exported files; page variables in the order they are selected on the Table tab are first, then row variables, column variables, and finally, the statistics/values.

### **Variable Recodes**

Lists the valid groupings for each table variable. When the Numeric Representation option is selected to format the variables as numbers, the variable groupings are listed with the numbers assigned to them. Each grouping is given a unique number starting at zero in the order the groupings are listed in the variable definition. The data dictionary lists the name of each table variable followed by a cross-walk list of the variable groupings and the corresponding number.

## **Export Window**

SEER\*Stat gives you the option to create a text-based data file containing all the statistics/values generated from your analysis. This data file can be used as input into other standard software packages. Any software package which can read text files can import the text-based output file. Using the options on the Export Window you can select the options which will make your data file the most compatible with your software package.

### **Data File**

Specify the path and name for your exported text file. The text file may be saved either as a standard ASCII file by using the “.txt” file extension or as a gzipped text file by using the “.gz” file extension. Actually, you may

specify any file extension, but only the “.gz” extension will result in a gzipped file.

- If you have not saved the matrix, the default path will be the same as the path for your user-defined variables listed on the Preferences Window and the name will be “export.txt”.
- If you have saved the matrix, the default path and name will be the same as the saved matrix path and name, except the file extension will be “.txt”.
- If you are uncertain about the exact path of the directory where you wish to save your exported text file, click the Browse button to select the path.

### **Dictionary File**

Specify the path and name for the text-based data dictionary. The output data dictionary gives information about the exported data text file. In particular, it records the path name for the text file and the path name for the original matrix, lists the variables and statistics/values included in the text file in the order of appearance (left to right), and the valid values for each variable with recodes when you have selected the Numeric Representation option.

- If you have not saved the matrix, the default path will be the same as the path for your user-defined variables listed on the Preferences Window and the name will be “export.dic”.
- If you have saved the matrix, the default path and name will be the same as the saved matrix path and name, except the file extension will be “.dic”.
- If you wish to save the dictionary in the same directory as the data file, follow these steps:
  - 1 Highlight the path and name of the data file and copy to Windows clipboard by pressing Cntrl+C.
  - 2 Click inside the Dictionary File text box. Paste the path and name of the data file from the Windows clipboard by pressing Cntrl+V.
  - 3 Change the name of the dictionary file so that it is unique. We suggest that you change the file extension from “.txt” or “.gz” to “.dic”.
- If you are uncertain about the exact path of the directory where you wish to save your exported text file, click the Browse button to select the path.

### **Output Variables as**

You may select to have the table variables coded as numbers or as the label text with or without quotes. The code equivalents will appear in the

dictionary file. Choose the coding method that will be the most compatible with your software package. The Numeric Representation option is disabled in the Case Listing Export Options.

**Numeric Representation** The groupings of each variable will be recoded to numbers starting with zero in the order the groupings are listed in the variable's definition. For example, if the variable were Sex with the values in order "Male and female", "Male", and "Female", then the recode would be 0 = "Male and female", 1 = "Male", and 2 = "Female". If the order of the variable groupings in the database dictionary is changed to "Male", "Female", and "Male and female", then the recode would be 0 = "Male", 1 = "Female", and 2 = "Male and female".

**Labels Enclosed in Quotes** The labels of the variable groupings/values will be the same as the text shown on your matrix, with double quotes surrounding them. For a Case Listing matrix, Unformatted labels will not appear in quotes in the data file.

**Labels without Quotes** The labels of the variable groupings/values will be the same as the text shown on your matrix, without quotes unless you have selected the Enclose Fields Containing Delimiter in Quotes option.

### **Line Delimiter**

The information found in the lines of the data file depends on the type of matrix you are exporting. For a Frequency/Rate matrix, each line contains the statistics calculated for a combination of the table variables. For a Survival or Case Listing matrix, each line is the same as a single row from a single page of the matrix. Lines are delimited differently by DOS/Windows and UNIX. Choose the option that matches the platform where you are saving your export text file.

**DOS/Windows (CR/LF)** DOS/Windows expects both a carriage return (CR) and a line feed (LF).

**UNIX (LF)** UNIX expects only a line feed (LF).

### **Field Delimiter**

The fields of the data file also depend on the type of matrix you are exporting. For a Frequency/Rate matrix, the fields will be the table variables and each statistic calculated (for example Rate, Count, and Population). For a Survival or Case Listing matrix, the fields are the same

as the columns on the page. You should select the delimiter that will be the most compatible with your software.

<b>Tab</b>	Places a single tab between each field in the export data file. Since tabs never appear in the text or numbers within any field, it was selected to be the default delimiter.
<b>Space</b>	Places a single space between each field in the export data file. If you choose to output the variables as labels, be careful to either enclose them in quotes or check the Enclose Fields Containing Delimiter in Quotes. You should also consider which Missing Character code to use, if applicable.
<b>Comma</b>	Places a single comma between each field in the export data file. If you choose to output the variables as labels which may have commas in them, be careful to either enclose them in quotes or check the Enclose Fields Containing Delimiter in Quotes. In addition, you may want to check the option to Remove All Thousands Separators from the numeric fields.
<b>Semi-Colon</b>	Places a single semi-colon between each field in the export data file. It is unlikely that a semi-colon would appear in the labels, so this option is desirable when your software does not easily deal with quotes and tabs.

### **Missing Character**

When a statistic cannot be calculated in a Frequency/Rate or Survival session, the cell of the matrix is marked with a footnote. No number appears in the cell. Choose a missing character to represent the empty cell. This option is disabled in a Case Listing Export Options.

<b>Space</b>	A space is placed wherever the value of a missing cell would appear. If you have chosen a space as the field delimiter, you may not want to choose a space as the missing character.
--------------	--

**Period**

A period is placed wherever the value of a missing cell would appear. You should consider if your software package will handle a period as a numeric missing rather than as text.

**Enclose Fields Containing Delimiter in Quotes**

This option will place double quotes around any field value which contains the delimiter. For example, if you selected the comma for your field delimiter, the label *All Sites* will not appear with quotes in your data file, but "*Uterus, NOS*" will because it contains a comma.

**Remove all Thousands Separators (Commas)**

This option removes all commas acting as thousands separators from the calculated statistics. You may wish to select this option, even if you are using a field delimiter other than the comma. The absence of the thousands separators may be easier for your software package to handle.

**Remove Flags (Footnote Characters)**

When checked, this option removes all footnote characters from the data file. Footnote characters can indicate such things as a positive significance test or the reason for an empty cell. When deciding whether to select this option, you should consider if the footnote information will be helpful in your analysis of the data file. This option is disabled in a Case Listing matrix.

**Output Variable Names Before Data**

This option lists the variable names before the data in the dictionary file. The variable names are always in quotes. This can be useful when you import/copy data into another software package that uses field names.

**Defaults**

Click this button if you make changes to the Export options and wish to return to the default settings.

**Set Default**

Once you have edited the Export options, you may set the current settings as the defaults by clicking this button. This may be useful if you use certain export settings regularly. These options will now be the defaults anytime you open a session or matrix in SEER\*Stat.

## **Output Variables as Numeric or Text in Export File**

On the Export Window, you may select to have the table variables coded as numbers, or as the label text with or without quotes. The code equivalents will appear in the dictionary file. Choose the coding method that will be the most compatible with your software package.

**Numeric Representation**

This option is only available in Frequency/Rate or Survival matrices. The groupings of each variable will be recoded to numbers starting with zero. For example, if the variable were Sex with the values in order “Male and female”, “Male”, and “Female”, then the recode would be 0 = “Male and female”, 1 = “Male”, and 2 = “Female”. If the order of the variable groupings in the database dictionary is changed to “Male”, “Female”, and “Male and female”, then the recode would be 0 = “Male”, 1 = “Female”, and 2 = “Male and female”.

**Labels Enclosed in Quotes**

The labels of the variable groupings will be the same as the text shown on your matrix, with double quotes surrounding them. For a Case Listing matrix, Unformatted labels will not appear in quotes in the data file.

**Labels without Quotes**

The labels of the variable groupings will be the same as the text shown on your matrix, without quotes unless you have selected the Enclose Fields Containing Delimiter in Quotes option.

**Line Delimiter in Export File**

The information found in the lines of the data file depends on the type of matrix you are exporting. For a Frequency/Rate matrix, each line contains the statistics calculated for a combination of the table variables. For a Survival or Case Listing matrix, each line is the same as a single row from a page of the matrix. Lines are delimited differently by DOS/Windows and UNIX. From the Export Window, choose the option that matches the platform where you are saving your export text file.

**DOS/Windows (CR/LF)**

DOS/Windows expects both a carriage return (CR) and a line feed (LF).

**UNIX (LF)**

UNIX expects only a line feed (LF).

**Field Delimiter in Export File**

The fields of the data file depend on the type of matrix you are exporting. For a Frequency/Rate matrix, the fields will be the table variables and each statistic calculated (for example Rate, Count, and Population). For a Survival or Case Listing matrix, the fields are the same as the columns on the page. You should select the delimiter that will be the most compatible with your software from the Export Window.



**Tab**

Places a single tab between each field in the export data file. Since tabs never appear in the text or numbers within any field, it was selected to be the default delimiter.

**Space**

Places a single space between each field in the export data file. If you choose to output the variables as labels, be careful to either enclose them in quotes or check the Enclose Fields Containing Delimiter in Quotes. You should also consider which Missing Character code to use, if applicable.

**Comma**

Places a single comma between each field in the export data file. If you choose to output the variables as labels which may have commas in them, be careful to either enclose them in quotes or check the Enclose Fields Containing Delimiter in Quotes. In addition, you may want to check the option to Remove All Thousands Separators from the numeric fields.

**Semi-Colon**

Places a single semi-colon between each field in the export data file. It is unlikely that a semi-colon would appear in the labels, so this option is desirable when your software does not easily deal with quotes and tabs.

## Missing Character in Export File

When a statistic cannot be calculated in a Frequency/Rate or Survival session, the cell of the matrix is marked with a footnote. No number appears in the cell. From the Export Window, choose a missing character to represent the empty cell. This option is disabled in the Case Listing Export options.

**Space**

A space is placed wherever the value of a missing cell would appear. If you have chosen a space as the field delimiter, you may not want to choose a space as the missing character.

**Period**

A period is placed wherever the value of a missing cell would appear. You should consider if your software package will handle a period as a numeric missing rather than as text.

## Enclose Fields Containing Delimiter in Quotes in Export File

On the Export Window, this option will place double quotes around any field value which contains the delimiter. For example, if you selected the

comma for your field delimiter, the label *All Sites* will not appear with quotes in your data file, but “*Uterus, NOS*” will because it contains a comma.

## **Remove all Thousands Separators (Commas) from Export File**

On the Export Window, this option removes all commas acting as thousands separators from the calculated statistics. You may wish to select this option, even if you are using a field delimiter other than the comma. The absence of the thousands separators may be easier for your software package to handle.

## **Remove Flags (Footnote Characters) from Export File**

On the Export Window of a Frequency/Rate or Survival matrix, this option removes all footnote characters from the data file. Footnote characters can indicate such things as a positive significance test or the reason for an empty cell. When deciding whether to select this option, you should consider if the footnote information will be helpful in your analysis of the data file.

## **Output Variable Names Before Data**

On the Export Window, if this option is checked, the names of the variables will be listed in the first line of the exported data file. The variable names will be displayed in quotes.

## **Edit Export Options**

Change the settings on the Export Window so that your exported data file will have the correct format. If you will be using another software package to load and process this data file, check that software package's requirements for importing text files.

### **Defaults**

Click this button to reset the Export options to the current set of default options.

### **Set Default**

Click Set Default to save the current settings as the defaults. This set of options will be restored whenever you click “Defaults” in the future. These will be the Export options each time you open the Export Window.



# Algorithms

## Frequency/Rate Algorithms

The algorithms involved with producing the statistics for the Frequency/Rate Sessions can be grouped into two main categories: algorithms for rates and trends.

### Note

The source code for the chi-squared and t distribution functions were obtained from the Library of C Routines for Cumulative Distribution Functions, Inverses, and Other Parameters (February 1994).

## Algorithms for Rates

### Crude Rate

A crude rate is the number of cases per 100,000 in a given population.

$$cruderate = \frac{count}{population} \times 100,000$$

### Age-Adjusted Rate

An age-adjusted rate is a weighted average of crude rates, where the crude rates are calculated for different age groups and the weights are the proportions of persons in the corresponding age groups of a standard population. Several sets of standard populations are included in SEER\*Stat. These include the total U.S. populations (1940, 1950, 1960, 1970, 1980, and 1990), an estimate of the U.S. 2000 population, 1991 Canadian population, and the world population. The age-adjusted rate for an age group comprised of the ages x through y is calculated using the following formula:

$$aarate_{x-y} = \sum_{i=x}^y \left[ \left( \frac{count_i}{pop_i} \right) \times 100,000 \times \left( \frac{stdmili_i}{\sum_{j=x}^y stdmili_j} \right) \right]$$

where count is the number of cases for the ith age group, popi is the relevant population for the same age group, and stdmili is the standard population for the same age group.

### Standard Error for a Crude Rate

This calculation assumes that the cancer counts have Poisson distributions.

$$SE_{crude} = \frac{\sqrt{count}}{population} \times 100,000$$

### Standard Error for an Age-Adjusted Rate

This calculation assumes that the cancer counts have Poisson distributions. Suppose that the age-adjusted rate is comprised of age groups x through y.

$$SE_{AARate} = \left[ \sum_{i=x}^y \left( \frac{stdmil_i}{\sum_{j=x}^y stdmil_j} \right)^2 \times \left( \frac{count_i}{population_i^2} \right) \right]^{1/2} \times 100,000$$

### Crude Rate Confidence Intervals

The endpoints of a p x 100% confidence interval are calculated as:

$$CI_{low} = \frac{\left( \frac{1}{2} \left( ChiInv \left( \frac{p}{2}, 2 \times count \right) \right) \right)}{population} \times 100,000$$

$$CI_{high} = \frac{\left( \frac{1}{2} \left( ChiInv \left( 1 - \frac{p}{2}, 2 \times (count + 1) \right) \right) \right)}{population} \times 100,000$$

where Chi Inv(p,n) is the inverse of the chi-squared distribution function evaluated at p and with n degrees of freedom, and we define Chi Inv (p,0) = 0.

Although the normal approximation may be used with the standard errors to obtain confidence intervals when the count is large, we use the above exact method that holds even with small counts (see Johnson and Kotz, 1969, or Fay and Feuer, 1997). When the count is large the 2 methods produce similar results.

### Age-Adjusted Rate Confidence Intervals

Suppose that the age-adjusted rate is comprised of age groups x through y, and let:

$$w_i = \frac{stdmil_i}{\left( pop_i \times \sum_{j=x}^y stdmil_j \right)}$$

$$w_m = \max(w_i)$$

$$v = \sum_{i=x}^y (w_i^2 \times count_i)$$

The endpoints of a  $p \times 100\%$  confidence interval are calculated as:

$$CI_{low} = \left( \frac{v}{2 \times rate} \right) \times \left( ChiInv \left( \frac{p}{2}, \frac{(2 \times rate^2)}{v} \right) \right) \times 100,000$$

$$CI_{high} = \left( \frac{v + w_m^2}{2(rate + w_m)} \right) \times \left( ChiInv \left( 1 - \frac{p}{2}, \frac{2(rate + w_m)^2}{(v + w_m^2)} \right) \right) \times 100,000$$

This method for calculating the confidence interval was developed in Fay and Feuer (1997). The method produces similar confidence limits to the standard normal approximation when the counts are large and the population being studied is similar to the standard population. In other cases, the above method is more likely to ensure proper coverage.

#### Note

The rate used in the above formulas is not per 100,000 population.

## Algorithms for Trends

### Percent Change (PC)

The percent change in rates over a particular time period is calculated by taking the difference of the first two years average rate and the last two years average rate. The difference is then divided by the average rate of the first two years and multiplied by 100 to convert it to a percent.

$$PC_{x-y} = \left( \frac{\left( (Rate_y + Rate_{y-1}) - (Rate_x + Rate_{x+1}) \right)}{(Rate_x + Rate_{x+1})} \right) \times 100$$

### Estimated Annual Percent Change (EAPC)

The Estimated Annual Percent Change (EAPC) is calculated by fitting a least squares regression line to the natural logarithm of the rates using the calendar year as a regressor variable.

$n$  = number of years

$r$  = rates

$y = Ln(r)$

$x$  = calendar year

$y = mx + b$

$$EAPC = 100 \times (e^m - 1)$$

Because the methods used in the calculation of PC and EAPC are not directly related, it is possible that the signs of the PC and the EAPC may disagree.

**Significance Test: EAPC to 0**

Testing the hypothesis that the EAPC is equal to zero is equivalent to testing the hypothesis that the regression parameter  $m$  is equal to zero. The hypothesis is rejected at a significance level  $p$  if  $\text{ProbT}(\text{abs}(m/\text{SE}_m), n-2) \geq 1-p/2$ , where  $\text{ProbT}(x, n)$  is the  $t$  distribution function evaluated at  $x$  and with  $n$  degrees of freedom, and where  $\text{SE}_m$  is the standard error of  $m$  from the regression.

**Standard Error for EAPC**

The standard error, i.e.,  $\text{SE}_m$ , is obtained from the fit of the regression (Kleinbaum, 1988). This calculation assumes that the rates increased or decreased at a constant rate over the entire calendar year interval. The validity of this assumption is not assessed. In those few instances where at least one of the rates is equal to zero, the linear regression is not calculated.

**Confidence Intervals for EAPC**

The endpoints of a  $p \times 100\%$  confidence interval are calculated as:

$$CI_{Low} = \left( e^{(m - (Tval * SE_m))} - 1 \right) \times 100$$

$$CI_{High} = \left( e^{(m + (Tval * SE_m))} - 1 \right) \times 100$$

Where  $Tval = T_{inv}(1-p/2, n-2)$  is the inverse of the  $t$  distribution function evaluated at  $1-p/2$  and with  $n-2$  degrees of freedom.

**Significance Test: EAPC to EAPC**

The differences between trends for two time periods are tested for statistical significance by comparing the difference in regression coefficients divided by the standard error of that difference with a  $T$  distribution with degrees of freedom defined as the sum of the years in both time periods minus 4 (Kleinbaum, 1988).

## Survival Algorithms

The algorithms involved with producing the survival statistics can be grouped into two main categories:

- Output rate table algorithms
- Special follow up interval algorithms

The Symbol Definitions for Survival Algorithms describe each symbol used in the survival algorithms. The survival algorithms are defined in the following sections.

## Symbol Definitions for Survival Algorithms

$X$	=	Interval number
$L_x$	=	Number of cases alive at the beginning of the interval
$D_x$	=	Number of cases who died during the interval
$U_x$	=	Number of cases untraced (lost to follow up) during the interval
$W_x$	=	Number of cases withdrawn alive during the interval
$L_x^*$	=	Adjusted number of cases alive during the interval
$P_x$	=	Observed survival rate for the interval
$CP_x$	=	Cumulative observed survival rate
$P_x^*$	=	Expected survival rate for the interval
$CP_x^*$	=	Cumulative expected survival rate
$R_x$	=	Relative survival rate for the interval
$CR_x$	=	Cumulative relative survival rate
$2SP_x$	=	2 standard errors for the observed interval survival rate
$2SCP_x$	=	2 standard errors for the observed cumulative survival rate
$2SR_x$	=	2 standard errors for the relative interval survival rate
$2SCR_x$	=	2 standard errors for the relative cumulative survival rate

## Algorithms for the Survival Output Rate Table

The relative survival algorithms of the output rate table depend on numbers obtained from the observed survival algorithms and the expected rate table. The observed survival algorithms are independent of both the relative survival algorithms and the expected rate table.

### Observed Survival Section of Output Rate Table

$L_1$	=	total number of cases included in the output rate table
$L_x$	=	number of cases alive at the beginning of interval x
$D_x$	=	number of cases dying during interval x.
$U_x$	=	number of cases untraced (lost to follow up) during interval x where the definition for "lost to follow up" is: alive with follow up date prior to some prescribed date.



$W_x$  = number of cases withdrawn alive during interval x where the definition of “withdrawn alive” is: alive with follow up date later than some specified date. (This is the same date as in previous definition.)

$$L_x = L_{x-1} - D_{x-1} - U_{x-1} - W_{x-1}$$

$$L_x^* = L_x - \frac{1}{2}(U_x + W_x)$$

$$P_x = 1 - \frac{D_x}{L_x^*}$$

$$CP_x = P_1 \cdot P_2 \cdot \dots \cdot P_x$$

### Relative Survival Section of Output Rate Table

$$P_x^* = \frac{1}{L_x} \sum_{i=1}^{L_x} \tilde{P}_i$$

where

- 1  $\tilde{P}_i$  = each individual's (i's) expected probability for surviving interval x.
- 2 The sum is over all individuals entering interval x alive.

$$CP_x^* = \frac{1}{L_1} \sum_{i=1}^{L_1} (\tilde{P}_1 \cdot \tilde{P}_2 \cdot \dots \cdot \tilde{P}_x)$$

where  $\tilde{P}_1, \tilde{P}_2, \dots, \tilde{P}_x$  refers to the expected probability of each individual (i) for surviving intervals 1, 2, ..., x and the sum is over all individuals in the table.

$$R_x = P_x / P_x^*$$

$$CR_x = CP_x / CP_x^*$$

$$2SR_x = 2R_x \sqrt{\frac{D_x}{L_x^* (L_x^* - D_x)}}$$

$$2SP_x = 2P_x \sqrt{\frac{D_x}{L_x^* (L_x^* - D_x)}}$$

$$2SCR_x = 2CR_x \sqrt{\sum_{j=1}^x \frac{D_j}{L_j^* (L_j^* - D_j)}}$$

$$2SCP_x = 2CP_x \sqrt{\sum_{j=1}^x \frac{D_j}{L_j^* (L_j^* - D_j)}}$$

## Algorithms for the Special Survival Follow up Interval

The special survival follow up intervals are sets of the regular intervals. Each special interval range has a beginning interval (a) and an ending interval (x). The special interval information is displayed at the bottom of each output rate table and can be divided into an Observed Survival and Relative Survival.

### Observed Survival

$$CP_{a-x} = P_a \cdot P_{a+1} \cdot \dots \cdot P_x$$

### Relative Survival

$$CP^*_{a-x} = \frac{1}{L_a} \sum_{i=1}^{L_a} (\tilde{P}_a \cdot \tilde{P}_{a+1} \cdot \dots \cdot \tilde{P}_x)$$

where the sum is over all persons entering interval “a” alive.

$$CR_{a-x} = CP_{a-x} / CP^*_{a-x}$$

$$SCR_{a-x} = CR_{a-x} \sqrt{\sum_{j=a}^x \frac{D_j}{L_j^* (L_j^* - D_j)}}$$

## Expected Survival Rates Table

The expected rate table which comes with SEER\*Stat contains the expected probabilities of survival generated from U.S. population statistics. Expected rates are obtained from the expected rate table by using each input case record's race, sex, age, and the age date (i.e., the date at which the age was coded).

- The Sex values available in the expected rate table are Male and Female only.
- The Age values available in the expected rate table are given in single years of age from 0 to 118.
- The Race and Year values available in the expected rate table depend on which expected rate table you use.

**U.S. 1970, 1980, 1990 (White, Black, Other)** matches to the cohort cases using the SEER Race Recode A variable and applies the 1970 rates to 1973-1975 ages, the 1980 rates to the 1976-1985 ages, and 1990 rates to the ages recorded in 1986 through the year of latest diagnosis.

**U.S. 1970, 1980 (Expanded Races)** matches to the cohort cases using the SEER Race Recode B variable and applies 1970 rates to the 1973-1975 ages and the 1980 rates to the ages recorded in 1976 and after.

This is a sample of an expected rates table for surviving 1-year for white, males, with ages 0 through 6 for 1970 and 1980.

**Sample Expected Survival Rate Table**

Race	Sex	Age	Age Date	
			1970	1980
White	Male	0	.98407	.98769
White	Male	1	.99901	.99908
White	Male	2	.99926	.99934
White	Male	3	.99941	.99947
White	Male	4	.99949	.99957
White	Male	5	.99954	.99961
White	Male	6	.99957	.99963

All rates saved in the Expected Rate Table are stored in a binary numerical format.

## Indexing the Expected Probabilities for Each Individual Patient

Let     $r$         =        race of individual  
           $s$         =        sex of individual  
           $g$         =        age at entry into life table  
           $y$         =        calendar year of entry into life table  
           $x$         =        interval of follow up

**Index for Expected Probabilities**

Annual Interval	Race	Sex	Age	Age Date
1	$r$	$s$	$g$	$y$
2	$r$	$s$	$g + 1$	$y + 1$
3	$r$	$s$	$g + 2$	$y + 2$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
$x$	$r$	$s$	$g + x - 1$	$y + x - 1$

For non-annual intervals, the expected rate table is expanded from years to months by taking the twelfth root of the yearly survival rate. For example, if the annual (January to December) expected survival rate for white males age 6 in 1970 is 0.99957, then 0.99996 would be the survival rates for each month of 1970.

# Bibliography

## Frequency/Rate Bibliography

Brown BW, Lovato J, and Russell K. 1994. *DCDFLIB: Library of C Routines for Cumulative Distribution Functions, Inverses, and Other Parameters*. University of Texas, M.D. Anderson Cancer Center. This work was supported by grant CA-16672 from the National Cancer Institute.

Fay MP and Feuer EJ. 1997. Confidence intervals for directly standardized rates: A method based on the gamma distribution. *Statistics in Medicine* 16:791-801.

Johnson NL and Kotz S. *Discrete Distributions*. Wiley: New York, 1969.

Kleinbaum DG, Kupper LL, and Muller KE. *Applied Regression Analysis and Other Multivariate Methods*. North Scituate, Massachusetts: Duxbury Press, 1988: 266-268.

## Survival Bibliography

Axtell L. 1969. Life table program: Observed and relative rates. *Methodologic Note No. 24, End Results Evaluation Section*, pp. 73-76. Bethesda, Maryland: National Cancer Institute.

Brown CC. 1983. The statistical comparison of relative survival rates. *Biometrics* 39:941-948.

Buckley JD. 1984. Additive and multiplicative models for relative survival rates. *Biometrics* 40:51-62.

Ederer F, Axtell LM, and Cutler SJ. 1961. The relative survival rate: A statistical methodology. In: *NCI Monograph No. 6*, pp. 101-121. Bethesda, Maryland: National Cancer Institute.

Esteve J, Benhamou E, Croasdale M, and Raymond L. 1990. Relative survival and the estimation of net survival: Elements for further discussion. *Stat. Med.* 9:529-538.

Gibberd RW and Hakulinen T. 1988. A relative survival analysis package. *The American Statistician* 42:230.

Hakulinen T. 1977. On long-term relative survival rates. *J. Chronic. Dis.* 30:431-443.

Hakulinen T. 1982. Cancer survival corrected for heterogeneity in patient withdrawal. *Biometrics* 38:933-942.

Hakulinen T and Abeywickrama K. 1985. A computer program package for relative survival analysis. *Computer. Programs. Biometrics* 19:197-207.

Hakulinen T and Tenkanen L. 1987. Regression analysis of relative survival rates. *Royal. Stat. Soc.* 36:309-317.

Hakulinen T, Tenkanen L, Abeywickrama K, and Paivarinta L. 1987. Testing equality of relative survival patterns based on aggregated data. *Biometrics* 43:313-325.

Shambaugh EM, Young JL, Zippin C, Lum D, Akers C, and Weiss M. SEER Program: Self Instruction Manual for Cancer Registrars: Book 7: Statistics and Epidemiology for Cancer Registries, NIH Pub no. 94-2766.

*SEER Program Code Manual*, National Cancer Institute, NIH Pub no. 94-3766.

# Index

.dic .....	84, 85, 86
.fmt .....	43, 44
.fmx .....	41, 42
.gz .....	84, 85, 86
.si .....	12, 13
.sim.....	67, 68
.sl .....	36
.slm.....	67
.ss.....	21, 23
.ssm.....	67, 68
.txt .....	84, 85, 86

## A

About SEER*Stat .....	7
Accessing the Number Filter .....	79
Accessing the Selection Window.....	58
accurate search statements .....	17, 25, 64
actively followed .....	25, 26
Add a "Not" to a Search Statement .....	62
Add a New Grouping to a Merged Variable.....	53
Add a New Grouping to a User-Defined Variable.....	48
Add a Page Variable to a Survival Session.....	29
Add a Search Line.....	63
Add a Table Variable to a Case Listing Session .....	38
Add a Table Variable to a Frequency/Rate Session.....	18
Add All button.....	48, 49, 52, 53, 54
Add All Groupings to a Merged Variable .....	53
Add All Merged Groupings Window.....	54
Add All Variable Values to a User-Defined Variable.....	49
Add button.....	29, 48, 52, 53
Add Parentheses to a Search Statement .....	62
Add Rest button .....	48, 49
Add Rest of the Variable Values to a User-Defined Variable .....	49
Adjust for Heterogeneity in Withdrawal .....	35
Adjust Increasing Relative Rates .....	35
Adjust Relative Rates over 1.0 .....	35
adjust text.....	74, 76
age .....	14, 15, 16, 17, 25, 65
Age at Diagnosis Selection for a Frequency/Rate Session .....	17
age date .....	28, 96, 97
age recode variable.....	14, 16, 17
Age Variable for a Frequency/Rate Session .....	16

age-adjusted rates .....	13, 14, 15, 16, 90
algorithms.....	90, 93
Algorithms for Rates.....	90
Algorithms for the Special Survival Follow up Interval.....	96
Algorithms for the Survival Output Rate Table .....	94
Algorithms for Trends.....	92
alignment.....	74, 76, 77
And button.....	57, 58, 61
arrange variables .....	18, 29, 38, 54, 75
ASCII.....	8, 84, 85
auto-size columns .....	73, 74, 77
Available Variables.....	54, 55
Avoid Inadvertent Exclusions .....	29
Avoid Variable Groupings that Overlap .....	65

## **B**

Begin Date .....	30
behavior .....	26, 65
bibliography.....	97
bold .....	78
Both button.....	38
Browse button .....	9, 40, 85
Building Search Statements with Search Lines.....	58

## **C**

Cancel a Job .....	67
Cancel button .....	43, 67
cancer data .....	8
case counts .....	19, 20, 33, 74
case distribution .....	20
case files .....	16, 17, 25, 37
case listing .....	68, 69, 76, 77
Case Listing Data Tab.....	37
Case Listing Matrix Navigation.....	75
Case Listing Matrix Options .....	75
Case Listing Matrix Window .....	74
Case Listing Output Tab.....	39
Case Listing Selection Tab.....	37
Case Listing Session Overview .....	36
Case Listing Table Tab .....	38
Case Selection for a Survival Session .....	28
cases per.....	9, 20, 21
category .....	18, 28, 38, 40, 43, 54, 65
center text .....	74, 76
Change Alignment of Values in a Matrix .....	76
Change button.....	9
Change Format of Variable Values .....	77

Change Matrix Font.....	78
character - missing.....	85
choose statistics.....	13, 23, 34
Clear button.....	17, 25, 37
client-server configuration .....	10, 66
Close button .....	67
column.....	69, 71, 73, 75, 76
headings.....	69, 71, 74, 83
hide .....	75, 76
percentages .....	19, 20
variables.....	18, 38, 76
width.....	73, 74, 77
Column button.....	18, 38
combined option.....	49, 52, 53, 54
compares - EAPC.....	16
complete execution .....	10, 66
confidence intervals - EAPC.....	15, 19, 21, 70, 90, 92
conjunctions .....	61, 63
Conjunctions for Search Statements .....	61
Contents Tab.....	11
Context-Sensitive Help.....	11
continuous values .....	59, 60
Converted values .....	74, 76, 77
Copy a Matrix to the Windows Clipboard .....	82
copy selected values.....	74
Create a Merged Variable .....	45
Create a Search Line .....	58
Create a User-Defined Variable .....	45
Create an Exact Search Statement.....	65
Create button .....	20, 40, 43, 45
Create Merged Variable Window.....	52
creation time.....	68
crude rates .....	13, 14, 90
Cumulative Summary .....	34, 35
cumulative summary table.....	22, 28, 33, 34, 35

## D

data dictionary.....	84, 85
data files.....	84, 85, 88, 90
Data Tab .....	13, 23, 37, 39, 40, 44, 68
database .....	9, 13, 23, 37
dictionary.....	41, 42, 44, 46
Database for a Case Listing Session .....	37
Database for a Frequency/Rate Session.....	14
Database for a Survival Session .....	23
Database Variables.....	43
date .....	30, 31



of diagnosis.....	30
Dates for a Survival Session .....	30
defaults.....	9, 15, 16, 21, 26, 32, 76, 77, 78
export options.....	84, 87, 90
definitions.....	24, 39, 40, 41, 43, 47, 57
grouping .....	52, 55, 56
user-defined variable.....	9, 83
Delete a Grouping from a Merged Variable.....	55
Delete a Search Line from a Search Statement .....	64
Delete an Existing Grouping from a User-Defined Variable .....	51
Delete button.....	19, 29, 40, 46, 48, 51, 52, 55
Delete User-Defined or Merged Variables.....	46
delimiter.....	86, 87, 88, 89
diagnostic confirmation.....	26
Dictionary .....	39
dictionary file .....	39, 40, 84, 85, 86, 90
Dictionary Window.....	40
directory .....	9, 41, 42, 68
disabled options .....	51, 85
display .....	11, 79
flags .....	70, 72
footnotes .....	70, 72
p-values.....	21
rates .....	9, 19, 20, 33, 34, 79
row numbers .....	74
titles.....	70, 72, 75
Display As .....	74, 77, 90
Display p-values for EAPC Significance Testing .....	21
Display Rates as Cases Per ____.....	21
Display Standard Errors and Confidence Intervals.....	21
Display Tables.....	34
Displayed Variables .....	75, 76, 78
distribution.....	19, 20
Dn button.....	18, 28, 38, 50, 56, 75
DOS/Windows.....	86, 88

## E

EAPC .....	7, 12, 13, 14, 15, 16, 19, 20, 21, 93
Edit.....	11
Edit "Not" in the Search Statement.....	64
Edit a Grouping Definition in a Merged Variable .....	55
Edit a Search Line.....	59
Edit a Search Line or Change the Conjunction .....	63
Edit button .....	12, 15, 17, 25, 28, 37, 52, 55, 57, 58, 63
Edit Export Options .....	90
Edit menu .....	18, 19, 29, 38, 39, 75, 82
Edit Merged Grouping Window .....	52, 57

edit merged variable.....	46
Edit Merged Variable Window .....	52
Edit Parentheses in a Search Statement.....	55, 63
Edit P-Values Used in a Frequency/Rate Session .....	15
Edit the Search Statement .....	63
edit user information.....	10
edit user-defined variable .....	45
Edit Variable Window .....	47
e-mail .....	10, 11, 66
empty pages.....	35
Enclose Fields Containing Delimiter in Quotes in Export File.....	89
End Date .....	30
error message .....	20
Estimated Annual Percent Change .....	7, 12, 13, 14, 15, 16, 19, 21, 92
exclude.....	17, 25, 26, 30, 37, 56, 57, 58
Exclude All Death Certificate and Autopsy .....	27
Exclude All Invalid Age Dates .....	28
Exclude All Second and Later Primaries .....	27
Exclude All Sex not Male or Female .....	27
Exclude All Unknown Ages .....	27
Exclude All Unknown Races .....	27
exclude cases .....	16, 25, 37, 52, 56
Execute a Session .....	66
execution options .....	10, 66
expanded races.....	25, 96
Expected Rate Table.....	25
Expected Rate Table Standard Survival Selections .....	27
expected survival rates.....	23, 24, 35
Expected Survival Rates Table .....	96
Export button.....	41, 42
export data .....	69, 71, 74, 84, 86, 87
dictionary.....	84
Export Window .....	85, 88, 89, 90
file .....	85, 88, 89
options.....	85, 90
Export Data Files.....	84
Export Results to Other Applications.....	84
Export User Variable Definitions Window.....	42
Export User Variables .....	42
Export Window .....	85
ExportEdit.....	85, 88
 <b>F</b>	
FAQs .....	11
field delimiter .....	84, 85, 89
Field Delimiter in Export File .....	88
field names .....	85, 90

file .....	84, 85, 86
extension.....	12, 21, 36, 44, 67, 85, 86
names .....	41, 42, 68, 85
size.....	68
File menu ..	9, 12, 13, 20, 21, 23, 34, 36, 37, 39, 40, 41, 42, 43, 67, 68, 70, 72, 83
fill character .....	80, 81
fill side .....	80, 81
filter .....	69, 72, 79, 80, 81, 82
Filter Window .....	79
Find button .....	18, 28, 29, 38, 41, 43, 45, 58, 59
Find Tab .....	11
Find Window .....	59
Finish Remote .....	66
Fixed Width Filter .....	81
Flag Cum Relative SE's Greater Than ____ %.....	35
flags .....	33, 69, 70, 85, 87
follow up .....	26, 36
font .....	9, 69, 71, 74, 78, 79
Font Sample.....	79
Font Size .....	78
Font Style .....	78
Font Type .....	78
Font Window .....	78
footnotes .....	19, 20, 33, 69, 70, 71, 72, 85, 89
Formatted.....	74, 76, 77
Frequency/Rate Algorithms .....	90
Frequency/Rate Bibliography .....	97
Frequency/Rate Data Tab .....	13
Frequency/Rate Matrix Navigation .....	70
Frequency/Rate Matrix Options.....	70
Frequency/Rate Matrix Window .....	69
Frequency/Rate Output Tab.....	19
Frequency/Rate Selection Tab.....	16, 17
Frequency/Rate Session .....	12, 13, 15, 40, 90
Frequency/Rate Session Overview .....	12
Frequency/Rate Table Tab.....	18

## **G**

Getting Help .....	11
Grouping Definition .....	52
groupings .....	44, 47, 48, 52, 54, 55, 57, 59
gzipped.....	84, 85

## **H**

hard disk space .....	8
Help button.....	11
Help Contents .....	11

Help menu .....	11
heterogeneity .....	34, 36
Hidden Variables .....	75, 76, 77
hide .....	70, 74, 75
columns .....	74, 75
EAPC rows .....	70
footnotes .....	70, 72
PC rows .....	70
rates .....	70
zero count rows .....	70
Hide Columns in a Matrix .....	76
Hide Statistics When Less Than ____ Cases .....	20
highlight all .....	47, 76
hyperlink .....	10, 14, 47

## I

Import button .....	40, 41
Import User Variable Definitions Window .....	41
Import User Variable Window .....	41
Import User Variables .....	41
incidence .....	10, 14, 23, 37
database .....	7, 43
Include "Other" Grouping in a Merged Variable .....	56
increasing relative rates .....	35
Index Tab .....	11
Indexing the Expected Probabilities for Each Individual Patient .....	97
individual option .....	49
Information Management Services .....	7, 11
internet .....	8, 47
intervals - survival .....	30, 31, 32, 33, 34
Intervals for a Survival Session .....	31
invalid age date .....	28
invalid Begin Date .....	30

## J

job .....	66, 67
cancel .....	67
Job menu .....	67
Job Priority .....	9, 67
Job Status Window .....	66

## L

labels .....	19, 52, 54, 55, 56, 71
labels enclosed .....	86, 88
Leading Zero Filter .....	81
life tables .....	24, 33, 34, 35
Line Delimiter in Export File .....	88

Link Search Lines With a Conjunction (And/ Or) .....	61
Linking Search Lines to Make a Search Statement.....	61
list box .....	73
Lock Feature for a Case Listing Matrix.....	76
Lock Feature for a Matrix .....	73
log in .....	10, 66
Lost to Follow Up Date .....	31
low case counts.....	35

## M

malignant.....	25, 26, 65
manipulate matrix .....	70, 73, 75
matrix .....	66, 67, 68, 70, 73, 75
footnotes .....	70, 72
Matrix menu .....	12, 21, 68, 69, 71, 73, 74, 78, 79, 84
options.....	70, 72, 75
Matrix Footnotes .....	72
Matrix Navigation .....	70, 73, 75
Matrix Properties Window .....	68
Matrix Window.....	69, 71, 74
memory .....	8
Merge button .....	40, 45, 51
Merged Grouping Window.....	57, 58, 59, 64
merged variable.....	18, 29, 40, 41, 42, 43, 44, 45, 46, 47, 51, 52, 53, 54, 56, 57, 58, 63, 65
create .....	18, 28, 34, 35, 40, 43, 44, 45, 52, 54, 56, 57, 58, 62, 64, 65
grouping .....	52, 53, 54, 56, 64
modify.....	46, 52
Merged Variable Window .....	52
Merged Variables .....	44
microscopically confirmed .....	25, 26
missing character .....	86, 87, 89
Missing Character in Export File .....	89
Modify a Merged Variable .....	46
Modify a User-Defined Variable .....	45
Modify button.....	40, 45, 46
Months Per Interval .....	32
mortality data.....	10
Move Up/Move Dn buttons.....	18, 28, 29, 38, 47, 52, 55, 56, 71, 75

## N

name .....	47, 52, 85, 90
National Cancer Institute .....	8
navigation.....	70, 73, 75
Negatives Filter .....	82
New Merged Grouping Window .....	52, 53
Not button.....	53, 58, 61, 62, 63, 64

notification of job completion .....	10, 66
Number Filter .....	79
Number of Intervals .....	32
Numeric Representation .....	85, 86, 88

## O

Observed and Relative Survival .....	24
Observed Survival Only .....	24
observed survival rates .....	23, 24, 33, 94
OK button .....	9, 15, 40, 41, 42, 45, 46, 53, 55, 58, 59, 61, 63, 71
Open a Case Listing Session .....	36
Open a Frequency/Rate Session .....	25
Open a Matrix .....	67
Open a Survival Session .....	21
Open the Dictionary .....	40
Open the Edit Variable Window .....	47
Open the Merged Variable Window .....	51
operator .....	57, 58, 59, 62
options - matrix .....	20, 34, 69, 70, 71, 72, 74, 75
export .....	84, 90
Or button .....	57, 61, 63
order .....	38, 69, 71, 74, 77
Order Labels Window .....	71
Order Search Lines in a Search Statement .....	64
Order the Groupings of a Merged Variable .....	56
Order the Groupings of a User-Defined Variable .....	50
Order the Page Variables in a Survival Session .....	29
Order the Table Variables in a Case Listing Session .....	38
Order the Table Variables in a Frequency/Rate Session .....	18
Other Case Variables Selection for a Frequency/Rate Session .....	17
Other grouping in a merged variable .....	52, 56
output .....	67
matrix .....	9, 18, 19, 33, 39, 56, 66, 68, 69, 71, 72, 74, 75, 83
methods .....	66, 82, 83, 84, 85
title .....	70, 72, 75
variables .....	84, 85, 86, 87, 90
Output Matrix .....	67
Output Tab .....	19, 33, 39
Output Variable Names Before Data .....	90
Output Variables as Numeric or Text in Export File .....	88
Overview .....	8, 12, 22, 36

## P

Page button .....	18
Page Through a Matrix .....	70, 73
page variables .....	18, 19, 28, 29, 34, 73
page/cell count .....	68, 69

Parameters for a Frequency/Rate Session .....	15, 31
Parameters Tab .....	30
Parenthesis button .....	61, 62
password.....	10, 66
path .....	85, 86
percent change - PC .....	7, 13, 14, 16, 70, 92
Percentages for Frequency/Rate Sessions .....	20
population.....	7, 8, 14, 15, 43, 65, 96
Precedence in a Search Statement.....	72
Preferences.....	9
Prefix Characters Filter.....	82
prefix/ suffix characters .....	80, 82
primaries .....	27
Primary Data Location.....	9
Print a Session or Matrix .....	83
print preview .....	9, 20, 34, 39, 70, 72, 75, 83
Print User-Defined Variable Definitions Option .....	83
priority of job.....	9, 67
privacy.....	10, 20
progress .....	66
properties - view matrix .....	68, 69, 71, 74
p-values.....	13, 15, 21, 92
P-Values for a Frequency/Rate Session .....	15

## Q

quotes .....	86, 87, 88, 89
--------------	----------------

## R

Race Sex Year Dx Registry County Selection for a Frequency/Rate Session ...	17
rates .....	14, 15, 21, 23, 70
rearrange variables .....	18, 38, 50, 54, 56, 71, 75, 77
records - view .....	36, 37, 74
relative survival .....	7, 23, 24, 25, 32, 33, 35, 94, 95, 96
Remote Menu.....	10
Remove a Page Variable from a Survival Session.....	29, 63
Remove a Table Variable from a Case Listing Session .....	39
Remove a Table Variable from a Frequency/Rate Session.....	19
Remove all Thousands Separators (Commas) from Export File .....	89
Remove button.....	18, 29, 38
Remove Flags (Footnote Characters) from Export File.....	89
remove parentheses .....	63
remove search lines .....	63
Rename a Grouping in a Merged Variable .....	55
Rename an Existing Grouping in a User-Defined Variable.....	50
Rename button.....	48, 50, 52, 53, 55
Reorder Column or Row Variables in a Frequency/Rate Matrix.....	71
Reorder Column Variables in a Matrix.....	77

reporting source .....	27
Resize Columns or Rows in a Matrix.....	73, 77
Right Mouse Button menu.....	69, 71, 73, 74, 77, 79
Rounding Filter.....	81
row - variables.....	18, 19, 20, 68, 70, 73
height .....	73, 77
numbers .....	74, 75

## S

save.....	13, 23, 37, 46, 47, 68
matrix .....	68
session.....	13, 23, 37
to dictionary.....	9, 48
user-defined variables.....	9, 41, 45, 46
Save a Case Listing Session.....	37
Save a Frequency/Rate Session.....	13
Save a Matrix .....	68
Save a Merged Variable to the Database Dictionary.....	46
Save a Survival Session.....	23
Save a User-Defined Variable to the Database Dictionary.....	46
search for variables.....	59
search line.....	57, 58, 59, 60, 61, 63
Search Statements.....	57
Secondary Data Location.....	9
SEER .....	7, 8
SEER Cancer Data .....	8
SEER*Prep .....	9, 10, 14, 23, 24, 25, 37
Select All button .....	42
Select Both Continuous Values and Groupings for a Search Line .....	60
select cases .....	16, 25, 37
Select Columns or Rows in a Matrix.....	75
Select Continuous Values for a Search Line .....	60
select directory .....	9
Select Groupings for a Search Line .....	60
Select None button.....	42
Select Only Actively Followed .....	26
Select Only Malignant Behavior .....	26
Select Only Microscopically Confirmed .....	26
Select Variables Systematically .....	65
Selected Variables .....	54, 55
Selection Tab .....	16, 17, 25, 30, 37, 57, 58, 63, 65
Selection Window.....	57
Selection(s) Box on the Selection Window .....	59
sequence number.....	27
server configuration.....	10, 66
session.....	12, 22, 36
execute.....	66



print .....	9, 83
Session menu .....	66
view .....	68
set defaults.....	9, 15, 85, 90
Set the Job Priority.....	67
settings.....	9, 15, 84, 87, 90
sex.....	17, 28
shading.....	9, 83
show columns.....	70, 73, 75, 76
significance level/testing.....	15, 16, 19, 21, 92
software package .....	82, 87, 90
Sort button.....	36, 38, 39
sort variables.....	38
Special Intervals.....	32
Standard Error.....	20, 21, 24, 35, 70, 91, 93
Standard Life Table.....	34
standard million .....	13, 14, 43, 65
Standard Million for Frequency/Rate Session .....	16
Standard Survival Selections .....	26, 27
Standard Variables.....	43
Statistic for a Frequency/Rate Session .....	14
statistics .....	13, 23, 24, 25, 67, 70, 72
Statistics - Frequency/Rate Sessions.....	7
Statistics - Survival Sessions .....	7
Statistics Requiring a Population File.....	14
Study Cutoff Date.....	31
subset.....	16, 17, 25, 26, 28, 37, 53, 57, 58, 65
Suffix Characters Filter.....	82
summary intervals .....	34
Suppress if Zero Filter .....	80
Suppress Pages With Less Than __ Cases .....	35
survival .....	22
intervals.....	31, 32
matrix .....	67, 69, 72
rates .....	31
time .....	30, 31
Survival Algorithms.....	93
Survival Bibliography.....	97
Survival Data Tab.....	23
Survival Matrix Navigation.....	73
Survival Matrix Options .....	72
Survival Matrix Window .....	71
Survival Output Tab .....	33
Survival Parameters Tab.....	30
Survival Rates Definitions .....	24
Survival Selection Tab.....	25

Survival Session Overview .....	22
Survival Table Tab .....	28
Symbol Definitions for Survival Algorithms .....	94

## T

tab .....	88
Table Tab .....	18, 28, 30, 38, 59, 65
table variable .....	18, 28, 38, 85, 88, 90
Technical Support .....	11
text file .....	84, 85, 88
Thousands Separator Filter .....	80
thousands separators .....	79, 80, 85, 87
Tips For Accurate Search Statements .....	64
Title for a Case Listing Session .....	39
Title for a Frequency/Rate Matrix .....	20
Title for a Survival Matrix .....	34
toolbar .....	12, 13, 21, 23, 36, 37, 40, 66, 67, 68, 73, 83
Trend Variable for a Frequency/Rate Session .....	16
trends .....	7, 10, 13, 14, 15, 16, 21, 93

## U

Unformatted .....	74, 76, 77
UNIX .....	86, 88
unknown .....	27
unlabeled values .....	47, 60
untraced .....	31
Up/Dn buttons .....	18, 28, 29, 38, 48, 52, 54, 56, 71, 75, 77
Update button .....	48, 50
Update the Definition of an Existing Grouping in a User-Defined Variable .....	50
Use Shading When Printing .....	83
user information .....	10, 66
User-Defined Variables .....	44
user-defined variables - print .....	83
User-Defined Variables Location .....	9
Using Your Own Data .....	10

## V

values .....	44, 47, 48, 50, 57, 59
variable .....	18, 29, 38, 39, 40, 41, 43, 44, 57, 65
definitions .....	9, 39, 52, 54, 83
grouping .....	52, 54, 57, 58, 59, 65
name .....	57, 58, 60, 84, 85, 86, 90
values .....	54, 74
Variable Name .....	58
Variable Recodes .....	85
View the Session Which Produced the Matrix .....	68
view variables .....	39, 40, 52, 83

vital status .....	30, 31
Vital Status for a Survival Session .....	33

## **W**

website .....	7, 10, 11, 66
weighted rates .....	14
What You Will Need .....	8
width .....	73, 77, 79, 80, 81
Windows clipboard .....	74, 82
withdrawal .....	22, 34, 35
Working With Variables .....	44
www.seer.cancer.gov .....	11

## **Y**

year .....	15, 16, 17, 18, 25, 31, 65, 96
------------	--------------------------------

## **Z**

zero .....	70, 79, 80, 81
Z-Test Interval .....	31, 32